


```
Db 969 CPGRW----- 973
Qy 887 CPGWGHLDATGAGEKAPSPWGSIRTAQAHAHWTPAAGSCSVSCGRLMELRFLC--MD 944
Db 974 -----RAGS-----WS-----QCSVTCGNGTOERFVLCRTAD 1000
Qy 945 SALRVPVQEEELCGLASKPGSREVQAVPCPARWQYKLAACSVSCGGRVVRILYCARAH 1004
Db 1001 DSFGVCEE-----RETAR-ICRLGPCP-----RN 1025
Qy 1005 GEDGEEILLDTQCGLPRPEP-----QEACSLPEPCPRWKVMSLGPSCASCGLGTTARR-- 1058
Db 1026 TSDSPKSVV--VQWLSRPDPNSPQETSSKRCQ-----GDKSVFCEVLSRYC 1074
Qy 1059 -----SVAC-----VOLDQGGQVEVDEAACAALVRPEASVPCLIAD 1094
Db 1075 SIPGYNKLCCKSCNPDNLTDVDRAPPSGKHNDIEE-----LMPTLSVPTLME 1125

RESULT 3
T00017
Gene ADAMTS-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00017
R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A:Reference number: Z14055; MUID:98110583; PMID:9441751
A:Accession: T00017
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-951 <KUN>
A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A:Experimental source: strain 129SVJ
C:Genetics:

Query Match 12.2%; Score 944.5; DB 2; Length 951;
Best Local Similarity 30.1%; Pred. No. 4.2e-52;
Matches 239; Conservative 103; Mismatches 334; Indels 117; Gaps 27;

Qy 81 GRPPSPGQRRQRRAGGILHLELLVAVGPDVFOAHOEDTERVYLTNLNIGAEILLRD 140
Db 224 GKPSGPG---SIRKRFVSSPRVETMLVADQSMADFHGSLGKHYLLTLFSAARFYKH 279
Qy 141 PSLGAQFRVHLKXVILTEPEGAPNITANLTSLSVCGWSQTNPEDDTPDGHADLVLY 200
Db 280 PSIRNSISLVVVKILVIEEQKGPEVTSNAALTRNFCNWKQHNPSDRDPEHYDTAIL 339
Qy 201 ITRFDELEPDPGNRQVRGVTQLGACSPWTSCLITETDGLGVTTIAHEIGHSGFLEHDA 260
Db 340 FTRQDL-CGSHTCDTLGMADVGVCDPSCSVIEDDGLQAAFTAHGLHVFVNMPHDDA 398
Qy 261 PG-----SCGPGSHVNAS-----DGAAPRAGLAWSPCSRRLQLLSLSSAGRARCVWDPPRP 311
Db 399. KHCASLNGVTGDSHLMASMLSLSDHSP-----WSPCSAYWVTSFLDNGHGCECLMDKPQ- 452
Qy 312 QPGSAGHP---PDAQGLYYSANEQCRVARGPKAVACTFAREHLDMQALSCHTDPLDQS 368
Db 453 -----NPIKPSDLFTGLYDANRQOFTFGEESKHPDA---ASTCTTLWCTGTSGLL 503
Qy 369 SCSRLVLPLLDGTGECGVEKXCSGRCSLYELTPIAA-VHGRWSSWGPRSPSCSCGGV 427
Db 504 VCQTKHFWADGTSCEGKVCVSGKCNKTKMFKHFAFPVHSGWGPWPGWDCSRTCGGGV 563
Qy 428 VTRRQCNPNRPAFGGRACVAGDLQAEKMCNTQAC-EKTLQEFMSQQCAITDQGPLRSPG 486
Db 564 QYTMRECDNPVKNKGKCEGKVRVRSNIEDCPDNGKTFREEQC-EAHNEFSKASFG 622
```

```
Qy 487 GASFYHGAAPVHSGQDALCRHMCRAIGBSFIMKRGDSFLDGTCTCMPSGSPREDGTLSCV 546
Db 523 NEPTVWTFPKAGVSPKDRCKLTCEAKGIGYFVLPQKVDGTPCSP-----DST-SVCV 676
Qy 547 SGSCRTFCDCGRMSQQVWDRCOVCGDNSTCSPRKSGFTAGRAREKVTTLTTPNLTSV 606
Db 677 QGQCVKAGCDRIIDSKKFKDKCGVCGNGSTCKKMSGIVTSTRP-GYHDIIVTIPAGATNI 735
Qy 607 YIAHH-----RPLFTHLAVRIG-GRYVYAGKMSISNTTYPSLLEDCRVEYRVALTEDRL 660
Db 736 EVKERNQSGRNRNGSFLAIRAADGTYLNG-----NFTLSTLEQDLTYKTVLRYSGSS 789
Qy 661 PRLEIRIWGLPQEDADIQVYRRYEGYCNLTPDITFTYFQPKPROA-----WV- 710
Db 790 AALERIRFSPLKEPTITQVLM-----VGHALRPKIFIFMKKKTESFNAIPTFSEWVI 844
Qy 711 --AAVVRPCSVSCAGLRWVNYSLDQARKELVETVQCO-----GSOQPPPAWPE 758
Db 845 EEW---GECSTKCSG-W-----VQCRDINGHPASECAKEVKPASTR 886
Qy 759 ACVLEPCPPYNAVGDFGSCSASCGGLRPRVRCVQAQSLKLTLPARCRAGAQQPAVA 818
Db 887 PCADLPC-PHWQVGWSPSCSKTCGKGYKRTLTKCVSHDGGVLSN----- 929
Qy 819 LETCNPPQPCPARW 831
Db 930 -ESCDPLKKPKHY 941

RESULT 4
T00355
Hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:Shikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403880; PMID:9734811
A:Accession: T00355
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <ISH>
A:Cross-references: EMBL:AB014586; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
A:Experimental source: brain
C:Genetics:

Query Match 9.6%; Score 746.5; DB 2; Length 837;
Best Local Similarity 30.8%; Pred. No. 1.4e-39;
Matches 211; Conservative 92; Mismatches 300; Indels 83; Gaps 25;

Qy 61 LQALEPQAVSSYLSPGA-----PLKGR-----PPSPGFORQORORRAAGGI 102
Db 162 LQPLEGGTFNSAGGPGCAHLIRKSPASGQGPCMNCVKAPLGPSP-----RPRAKEFASLS 217
Qy 103 LHLELLVAVGPDVFOAHOEDTERVYLTNLNIGAEILLDPSLGAQFRVHLKXVILTEPEG 162
Db 218 RVEVTLVADDKMAAFHGAFLKRYLLTVMAAAAKAFKPSIRNPVSLVTVTLVILSGEE 277
Qy 163 ANITANLTSSLLSVCGWSQTNPEDDTPDGHADLVLYITRFDLELDPGNRQVRGVTOLG 222
Db 278 GPQVGSAAQTILRSFCAWORGLNTPEDSDPHDFTAILFTRQDL-CGVSTCDTTLGMADV 336
Qy 223 GACSPWTSCLITETDGLGVTTIAHEIGHSGFLEHDAAGSGC---GP---SGHWASD 275
Db 337 TVCDPARSCAIVEDDGLQSAFTAAHELGHVFNMLHNS---KPCISLNGPLSTSRHWAPV 394
Qy 276 GAAPRAGLAWSPCSRRLQLLSLSSAGRARCVWDPPRQPGSAGHPDPAQPGYYSANEOCR 335
```

```
Db 395 MAHVDEEPMSPGSAFITDPLNGYGHCLLD----KPEAPLHLFVTPFGKYDADRCQ 450
QY 336 VAFGPKAVACTFAREHLDMCOALSCHTDPLDQSSCSRLVLLPLLDGTEGVEKWSKGRCR 395
Db 451 LTFEGDSRHCP---QLPFPAAALWCSHLNGHAMCQTKHSWADGTPCGPAACMGGRCL 507
QY 396 SILVEL-----TPIAAVHGRWSSWSPRSCSRSCGGVVTTRRQCNPRPAFGACRACVADL 451
Db 508 HMDQLQDFNIPOA---GGWGPWGPWGDSCRTCCGGVQFSSRDCRTPVRNGGKCEGRRT 564
QY 452 QAEMCNTOACEK-TQLEFMSQOQA-----RTDQPLRSPPGASFY-HWGAAPVHSGDAL 505
Db 565 RFRSNTEDCPTGASALTFRBEOCAANYHRTD--LFKSPFGPMDVPRYTGVAPODQCKLT 622
QY 506 CRHMCRAIGESFIMKRGDSFLDGTROMPSGPREDGTLSLCVSGSCRTFGCDGRDMSQVW 565
Db 623 CQ--ARALGVYVIE--PRVVDGTGTPCSPD-----SSVCVQGRCIHAGCDDRIIGSKKX 672
QY 566 DRCCVCGGDNSTCSPRKGSFTAGAREYVFTLVTPLNTSVYI-----ANRPIFTHLAV 620
Db 673 DKCMVCGGDSGCSKQSGSFRKER-YGYNVNVITIPAGATHILVRQGNFGRHSYI--LAL 729
QY 621 RI-GRYVYVAGKMSISPNNTTYPSELLEDGRVEYRVALTEDRLPRLEIRINWGPLOEDADI 679
Db 730 KLPDGSYALNGEYTLMPST--DVVLPGAVSLRYS---GATASETLSGHGFLAQPLTLQ 784
QY 680 VYRYGEEYGNLTRPDITFTYFQPKP 705
Db 785 VL-----VAGNPQDTRLRYSFFVRP 805

RESULT 5
T47158
hypothetical protein DKFzp762C1110.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47158
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: EMBL:AL162080
A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762C1110
C:Genetics:
A:Note: DKFzp762C1110.1

Query Match 8.8%; Score 679; DB 2; Length 550;
Best Local Similarity 30.4%; Pred. No. 1.6e-35;
Matches 185; Conservative 75; Mismatches 231; Indels 118; Gaps 27;

QY 263 SCGSPSGHYMAS-----DGAAPRAGLAWSPCSRRLSLISAGARCVWDPPRPPQGSAG 317
Db 4 NGVNDQSHWASMLSLNDSQP-----WSPCSAYMITSFLDNGHGELMDKQP----- 51
QY 318 HP---PDAQPLYSANEQCRVAFGPKAVACTFAREHLDMCOALSCHTDPLDQSSCSRL 374
Db 52 NPILQPLGDLPGTSYDANRQOQFTFGEDSKHCPDA---ASTCSTLWCTGTSGVLVQTKH 108
QY 375 VELLDTGTEGVEKWSKGRCSRSLVEL-----TPIAAVHGRWSSWSPRSCSRSCGGVVT 430
Db 109 PFWADGTSCEGKWCINGKCVNKTDRKHFTPTF---HSGNMGWGPWGDSCRTCCGGVQYT 165
QY 431 RQCNPRPAFGACRACVADLQAEKNTQAC-EKTQLEFMSQOQARTDQPLRSPGAS 489
Db 166 MRECNVPVKNKGKCEGKRVYRSCNLEDCPNNGKTFREEQC-EAHNEPSKASFGSGP 224
QY 490 FYHWGNAVHPSQGDA---LCRHMCRAIGESFIMKRGDSFLDGTFCMSPGREDGTLSCV 546
Db 225 AVEW---IPKYAGVSPKRCKLICQAKIGYFFVLQPKVVDGTPCSP-----DST-SYCV 275
```

```
QY 547 SGSCRTFGCDGRDMSQVMDRCQVCGGDNSTCSPRKGSFTAGRAREYVTFLTVTPTNLTSV 606
Db 276 QGOCVAGCDRIIDSKKKPKCGVCGGNGSTCKKISGVS TSAKP-GYHDIITIPGATNI 334
QY 607 YIANH-----RPLFTHLAVRIG-GRYVYVAGKMS-SPNNTTYPSELLEDGRVEYRVALTEDRL 660
Db 335 EVKQRNQRGSRNNGSFLAIKAADGTIILNGDYTLS-----TLEODIMVKGVLRYSGSS 388
QY 661 PRLEETRINWGPLOEDADIQVYRRYGEEYGNLTRPDITFTYFQPKPRQ-----AWY- 710
Db 389 AALERIRSPLEKEPTIQLV-----TVGNALRPIKIYTYFVFKKESFNAIPTTSAWY 443
QY 711 --WAAVRGPCSVSCGAGLRWVNYSCLDQAKELVETVQCQSQOP-----PAWPEA 759
Db 444 BEW-----GECSSKCELG--W-----QRLVECDING--QDASECAKEVKFASTRP 486
QY 760 CVLEPCPPYWAVDGFGPCSASCGGGLRERPRVRCVEAQSLLKTLPPARCAGAAQPAVAL 819
Db 487 CADHPC-PQWLGEWSSCKTCGKYKRSILKLSHDGVL-----SH 528
QY 820 ETCNPOPCP 828
Db 529 ESCDPLKCP 537

RESULT 6
C89114
protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biol
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C_e
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC25867.1; PID:g32394501; GSPDB:GN00023; CESP:C37C3
C:Genetics:
A:Gene: C37C3.6a
A:Map position: 5

Query Match 8.5%; Score 659.5; DB 2; Length 1558;
Best Local Similarity 24.3%; Pred. No. 9.1e-34;
Matches 239; Conservative 106; Mismatches 379; Indels 237; Gaps 45;

QY 408 GRWSSNGPRSPCSRSCGGVVTTRRQCNPRPAFGGRACVADLQAEKNTQACEKTOLE 467
Db 77 GNWGPVVPNECSRSCGGVQLEKRCQSGD-----CTGASVYISCNLNACE-SGTD 127
QY 468 FMSQOQARTDQPLRSPGASFYHWGAAPVHSGDALCRHMCRAIGESFIMKRGDSFLD 527
Db 128 FRAEQCSKFNDEAL-----DGNVHKW---TPY-KGNKNCCLVCKPESGNFYVWKADKVD 178
QY 528 GTRCMPSPGREDGTLSCVSGSCRTFGCDGRDMSQVMDRCQVCGGDNSTCSPRKGSF-T 586
Db 179 GTKC-----DSKSNDCIVDGECLPVGCDKJGSSLKFDKCGKCDGDSCTKTIEGRFDE 232
QY 587 AGRAREYVTFLTVTPLNTSVYIANRHLRPLFTHLAVRIGG-RYVYVAGKMSISPNNTTYPSEL 645
Db 233 RNLSPGYHIIILPEGATNIKIOEARKSTNWLAKNGSDHFLNGNLII-----QVEK 285
QY 646 DGRVEYRVALTEDRLPRLEIRINWGPLOEDADIQVYRRYGEEYGNLTRPDITFTYFQPKP 705
Db 286 EVEVGGTIFVYDDAEP--ETLSAQGPLSELTVALLEFKGSR-----DIAIKYEFSIP 336
QY 706 ROAWV-----WAAVRGPCSVSCGAGLRWVNYSCLDQAKELVETVQC-QGSOQPPAW 756
Db 337 LBEEDVMTYKFDNWT----PCSVSCGKGVQTRNLYCIDGKNKGRVDDDLCENNATKPPF 392
```


QY	757	PEACVLEBPCPPYVAVDGFPCASCGG-GLRERPRVC--VEAQSLIKTLPPARCRAGA	813
Db	393	EKSCETVDCAEWFTGDMWSSCSSTGQGOQYRVYCHQVFANGRRV-TVEDGNCT--VE	449
QY	814	OPVALETNCPQPCPARWEVBSPTTSAGAGLALENETCPGADGLEAPVTEG--PG	870
Db	450	RPVK-QTCNRFCAP-EMOAGPWSACEKCGDAFOYRSVTCRSEKEEGEGLAADACPA	507
QY	871	SVDEKLPAPE-----PCVGMSCPPGWHLDATSAKEKAPSPWGSIRTAQAHHVMTPAAG	925
Db	508	DBQEKFDERTCNLGPCEGLTFTVGEWNL-----SCATQPPCEYEWTVSEW	536
QY	926	SCSVSCGRLMELRFLCMDSALRVPVQBELCGLASKEGSRREVCOAVPCPARWQYKLAAC	985
Db	537	-----CTR-----CNDT-----EETREVTCKDSQGR-----AYPLEKC	564
QY	986	SVSCGRGVVRRILYCARAHGEDDGEHILLDTQCQGLPRPEQACSLLE-PCPPRWKVM	1044
Db	565	LV-----DNSTEPTDTR-----SCATQPPCEYEWTVSEW	594
QY	1045	GPCSASCGLTARRSVACVQLDQGDVEVDEAACAALVRPEASVPCLIAD-CTYRWHVGT	1103
Db	595	SKCTTECGHGHKTRRIVCAIHQNGGLEVDDEGHCOA-EKPEGKTNCTNEEKCTGTWYTS	653
QY	1104	WMSCSVSCGDIQRRDTCILGPOAQAPVADFCQHLKPVTVRGCVGQGTSPSLVP	1163
Db	654	WSECTAECGGSGDRVAVCLN-YDKKVPV-EWCDEAVKPSKQDCNVDDC-----PTCVD	706
QY	1164	HBEAAPGRTT-----ATPAGACGRQHLEPTGTIDMRGPGQADCAVAIGRP	1209
Db	707	SEFGCCPDNSTFATGEFNGFCNSCSETEFGCCADNVTVAT-----GPNKCEEFVESP	760
QY	1210	LGEVTVLRVLESSL-NCSEA---GDMLLIWLGRITWRKMKRKLDMTFSSKTNTLVVRQCG	1265
Db	761	LN-----LEADVADAEASGD-----APELCSVT---NENGEAVDVECA	797
QY	1266	R-----PGGVLLRYGSQLAPETFFRECDMQLFGPWGEIVSPSLSPATSNAG--GCRL	1316
Db	798	TIAPITALLGDGELIGNDTASNETIH--CSKTEFG-----CCPDWYTAASGKNEGCP	850
QY	1317	F 1317	
Db	851	F 851	

RESULT 7
T34395
Hypothetical protein C37C3.6b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
C:Accession: T34395; T34394
R:Geisel, C.; Bradshaw, H.
submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C37C3.
A:Reference number: 221518
A:Accession: T34395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2167 <GEI>
A:Cross-references: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b
A:Experimental source: strain Bristol N2; clone C37C3
A:Accession: T34394
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1555, 'SKP', <GR2>
A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
A:Experimental source: strain Bristol N2; clone C37C3
C:Genetics:
A:Gene: CESP:C37C3.6b; CESP:C37C3.6a
A:Map position: 5
A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/1

Query Match	8.5%	Score 659.5;	DB 2;	Length 2167;
Best Local Similarity	24.9%	Pred. No. 1.3e-33;		
Matches 239;	Conservative 106;	Mismatches 379;	Indels 237;	Gaps 45;
QY	408	GRWSSNGRPSRSCSGGVVTRRRQCNPRPAPFGGRACVAGADLOAEMCNTOACEKTQLE	467	
DB	77	GNWGPVWVNECSRSCSGGVQLEKRCQSGD-----CTGASVRYISCNLNACE--SGTD	127	
QY	468	FMSQCAQRTDQQLRRSPGGSFYHWAAGVPHSGDGLCRHMCRAIGSFIMKRGDSFLD	527	
DB	128	FRAEQCKFNDEAL-----DGNYHKW---TPY-KGKNCCELVCKPESGFFYKADKVV	178	
QY	528	GTRCMPSGREDGTLISLCVSGSCRTFCGDMDSQQVMDRCQVCGDGNSTCSPRKSP-T	586	
DB	179	GTKC-----DSKNDICVDGECPLVGCCKGLSSLPKCKGCKDGDGSTCKTIEGRFDE	232	
QY	587	AGRAREYVFLTVTPNLTSVYIANHRLPFLTHAVRIGG-RYVAVGKMSIPNTTYP	645	
DB	233	RNLSPGYHDIKILPEGATNIKIQAERKSTNNLALKNGSDHFLVINGNLI-----QVEK	285	
QY	646	DGRVEYRVALTEDRLPRLEIRIWPLOEDADIOVYRRYGEYGNLTPDITFTYQPKP	705	
DB	286	EVEVGCTIFVYDDAEP--ETLSAQGPLSEELTVALLFRKGR-----DTAKYFESIP	336	
QY	706	ROAWV-----WAAVRGPCSVSCGAGLRWNTYSCLDQARKELVETVQC-QGSQOPPAW	756	
DB	337	LEEEVDYMKFDNWT---PCSVSCGKVQTRNLYCIDGKNKGRVEDDLCEENNATKPEF	392	
QY	757	PEACVLEBPCPPYVAVDGFPCASCGG-GLRERPRVC--VEAQGSLIKTLPPARCRAGA	813	
DB	393	EKSCETVDCEAEWFTGDMWSSCSSTGQGOQYRVYCHQVFANGREV-TVEDGNCT--VE	449	
QY	814	OPVALETNCPQPCPARWEVSEPSCTSGAGLAGLALENETCPGADGLEAPVTEG---PG	870	
DB	450	RPVK-QTCNRFCAP-EMOAGPWSACEKCGDAFOYRSVTCRSEKEGESEKLLAADACPA	507	
QY	871	SVDEKLPAPE-----PCVGMSCPPGWHLDATSAKEKAPSPWGSIRTAQAHHVMTPAAG	925	
DB	508	DEQEKFDERTCNLGPCEGLTFTVGEWNL-----SCATQPPCEYEWTVSEW	536	
QY	926	SCSVSCGRLMELRFLCMDSALRVPVQBELCGLASKEGSRREVCOAVPCPARWQYKLAAC	985	
DB	537	-----CTR-----CNDT-----BETREVTCKDSQGR-----AYPLEKC	564	
QY	986	SVSCGRGVVRRILYCARAHGEDDGEHILLDTQCQGLPRPEQACSLLE-PCPPRWKVM	1044	
DB	565	LV-----DNSTEPTDTR-----SCATQPPCEYEWTVSEW	594	
QY	1045	GPCSASCGLTARRSVACVQLDQGDVEVDEAACAALVRPEASVPCLLIAD-CTYRWHVGT	1103	
DB	595	SKCTTECHGHKTRVICAIHONGGLEVVDEGHCOA-EKPEGKTNCTNEEKCTGTWYTS	653	
QY	1104	WMESVSCGDIQRRRDTCILGPOAQAPVADFCQHLKPVTVRGCVGQGTSPSLVP	1163	
DB	654	WSECTAECGGSGQDRVAVCLN-YDKKVPV-EWCDEAVKPSKQDCNVDDC-----PTCVD	706	
QY	1164	HEEAAAPORTT-----ATPAGACGRQHLEPTGTIDMRGPGQADCAVAIGRP	1209	
DB	707	SEFGCCPDNSTFATGEFNGFCNSCSETEFGCCADNVTVAT-----GPNKSGCEEFVESP	760	
QY	1210	LGEVTVLRVLESSL-NCSEA---GDMLLIWGLRITWRKMKRKLDMTFSSKTNTLVVRQCG	1265	
DB	761	LN-----LEADVADAEASGD-----APELCSVT---NENGEADVVECA	797	
QY	1266	R-----PGGVLLRYGSQLAPETFFRECDMQLFGPWGEIVSPSLSPATSNAG--GCRL	1316	
DB	798	TIAPITALLGDGELIGNDTASNETIH--CSKTEFG-----CCPDWYTAASGKNEGCP	850	
QY	1317	F 1317		
DB	851	F 851		

RESULT 8

Tl18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: Tl18856; T24653
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: Tl18856
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL>
A:A/Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone C02B4
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19917
A:Accession: T24653
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WI2>
A:A/Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone T07C5
C:Genetics:
A:Gene: CESP:C02B4.1
A:Map position: X
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 568/3

Query Match 8.5%; Score 657; DB 2; Length 1444;
Best Local Similarity 23.1%; Pred. No. 1.2e-33;
Matches 324; Conservative 137; Mismatches 508; Indels 434; Gaps 71;

Qy 105 LELLVAVGPDVFAHQEDTERYYVLTNINIGAEILLRDPESLGAFQRVHLVKVMILTEPEGAP 164
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 245 LEFSLALNNHVLYQQDTLTPLNDIVIVRYEMRTOPSSALSTGVHK----- 291

Qy 165 NITANLTSSLVSGVWSQINPEDD-TDPGHADLVLYITREDLELPDGNRQVRGYTLGG 223
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 292 --NGQAQLSDAFRCYQAHNPOTDLTMHYDHGVLLTGYDIY--HTTTSVAGVAPVAR 347

Qy 224 ACSPTWSCUITEDTGFDLGTVIAHEIGHISFGLEDHGA PGSCGPSGHVWASDGAAPRAGL 283
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 348 MCDPLFACSLVEGLHGSRFLVAHEMNHWMDHGVQNC-KNGKCCMLSAVNGAKT--- 404

Qy 284 AWSPCSRRQ----LLSLLSAGRARCVMDDPRPQPFG--SAGHPDDAQ-PGLYYSANEQCVRV 336
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 405 TWSDCSRVREFNAFLQLQDESGRGNCLD--ASPGLISTNHLSLDRLPGQRTAQOCSY 461

Qy 337 AFG-----PKAVACTFAREHLMCOALSCHTDPLDOSSCRLLVPLLDGTGECGVKWC 389
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 462 FNGRDYKVBIPIKGAMD-----DICRIWCG----NSGSTIATAHPALEGSWGCANKWC 511

Qy 390 SKGRGRS-LVELTPIAVHGRWSWG-----PRSPCSRSCGGVVTRRRQCNNRPARG 442
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 512 HRGQCCTHTFTGLTPV-PIDGEWSWGGAEKCPIQQCAVSGSITVQGHRDCVNPA PNG 570

Qy 443 GRACYGADIQAEMC--NTQACEKTQLEMSQQCATDGPPLRSSP--GGAGFYHWGA AV 497
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 571 GKTCEGANIRGI VCGATSSNCLGFTREBFNGKICSSIKYDKPKPOQLTGEGFEH--STQ 628

Qy 498 PHSQGDALCRMCAIGESFIMKRGDSDLGTRCMPSGFREDGTLSLCVSGSCRTFGCDG 557
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 629 P-----CRVWCHLTGSELIRNKG-QFPDTPC-----GFPAYCVGGQCLALSCDN 672

Qy 558 -----RWDSQQV-----WDRCQV-CGGDNSTCSPRKGSFTAGRARXVTFL 597
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 673 KALVEQPCDPRIEGRSVHGWEESSWSECVSQ----- 707

Qy 598 TVTPNLTSVIANHRLPLFTHLAIVIGRYVAVGKMSISPNTYPSSLLEDGRVEYRVALTE 657
 ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db 708 -----LGRVERERKCS-----SGRKCQGV--GE 729

Qy	658	DRLP-----PLEIRIWGFLQEDADIQVRYVGYEEVGNLTRDIT-----FTYFQPK 704
Db	730	ESRPECBGLRDCBEFGKEWKGSCCKALGVQKRP-----RCLTDQCSKSHLOEE 781
Qy	705	-----PRQAW-----WAAVRGCSVSCGAGLRWVNTSCLDQARKELVETVOCQSQPPAW 756
Db	782	RPCDNEGCWNTDWESS-----CSQSCGGRRYRIRKCLDD-----KCDGDDLE--- 825
Qy	757	PEACVLEPC-PPYNAVGFPCSBASCGGLRBRPVRVCVEAQGSLLKTLPPARCBAQAQP 815
Db	826	KESCNTOKCISQSW--GDWLPCSVSCGIGFQIRELJC--DGEL-----CATANKQ-- 871
Qy	816	AVALETNCPQPCPARWEYS-----EPSSCTSAGGAGLALENCTVFGADGLEAPVVTBGP 869
Db	872	---ARTNQOQCPGAFSUSVSEWGEWTTCSATCGEGLSRSECRGS-----CTHDD 922
Qy	870	GSDVDEKLPAPEPCVGMSCPPGWHLDATSAGBKAPSPWGSIRTGAQAHHVWTPAAGSCSV 929
Db	923	ASQTRR-----CVNGPC-----EHSYLTWSE-----WT-TCETCSS 952
Qy	930	SCGR-----GLMELRFLCMDSALRVPVOBELCGLASKPGSBRVCOAVPCPARWQVKL 982
Db	953	FDSKRRIAKCDGTTB--NCODK-----IDBETCDIA-----CLUREKHSFGSISPR-RPKL 999
Qy	983	AACSVSCRGVYRRILYCARAHGEDDEGEILLDTQCGLPRPQPQEAACSLPCCPPRWKVM 1042
Db	1000	IT-SNDLRKARGRPILPIESIHSE-----KWS--- 1025
Qy	1043	SLGPCSASCGLGTARRSVACVOLDOQGVDEFAACAALVRPRAV-----PCLIIDCTVRW 1099
Db	1026	EWGPCSVTCGSGRRVTRGC-----QEASC-----PEOHIQTEECNLSCLFLF 1069
Qy	1100	HVGTWMECSVCG-DGIQERRRCTLGPOAQAPYPAOF--COHLPKPVTYVRCWGAGPCVGQ 1156
Db	1070	IWSOWSSCSKCGQDGIQTRQKLCLPNNACSSYAESRRCKDLPS-----CSSI 1118
Qy	1157	GTPSLVPHHEAAAPKRTATPAGACG-----RQHLEPTGTIDMRGPGQADCAVA 1205
Db	1119	SSGRTISENGFADPRAWSEWSSACSFSLTSTRERFCQVVDPT-----VOGFCAGA 1170
Qy	1206	IGRPLGEVTVLVRLESSLNCAGDMLLLWGRLTWKMKRKLDDMTFSKTLVLVVRQCG 1265
Db	1171	I-----LEQIPCAPGSCSPAGGW-SLMSE--WSSCSKOCGD-----TGHQIRNRMCS 1215
Qy	1266	RP---GGVLLRYGSQLAPEIFYREC-DMQLFGPWGEIVSPSLSPATSNAGGCLFINV 1320
Db	1216	EPFSPNRGAYCGSYFQDQPCVMDNVCSDKVKDGGW----- 1251
Qy	1321	APHARIAIHALATNMCAGTEGANASYILIRDTSHLRTTAFHGQOVLXWSESSQAEFMS 1380
Db	1252	-----TDWTAWSECTD---YCRNGHRSRTRFCANPKPSOGGAQCTGSDFELN 1295
Qy	1381	BGFLKAQASLRQOYW-TLOSWPV 1402
Db	1296	PCFDPARCHLRDGGWSTGSDWTP 1318

RESULT 9

T22545
hypothetical protein F53B6.2 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22545
R:White, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19578
A:Accession: T22545
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1059 <WIL>
A:Cross-references: EMBL:Z81086; PIDN:CA803121.1; GSPDB:GN00019; CESP:F53B6.2
A:Experimental source: clone F53B6
C:Genetics:

A;Gene: CESP:F53B6.2
A;Map position: 1
A;Introns: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 84/3

Query Match 7.3%; Score 568; DB 2; Length 1059;
Best Local Similarity 23.3%; Pred. No. 3.8e-28;
Matches 195; Conservative 81; Mismatches 292; Indels 248; Gaps 32;

QY 403 IAAVHG---RWSSWGPRSPCRSRRGGVVRTRRRQCNPNRPAFGGRACVGAOLQAEMCNTQ 459
DB 16 IASFHVDLSWAANAFWSSCTKTCGGGVSHQLRRCLTSK-----CSGESVRFVKVCAQK 68
QY 460 ACEXTQLEFMSQQCARTDQPLRSSPGGSAFYHMGAAVPHSQGDALCHMCAIGESFIM 519
DB 69 TCESKSLRARDTIC-----GGBEI-----VSRGQCEVVCR--SLTIGANFLW 108
QY 520 KRGSFLLGTRCMPGPRDGLTSLCVSGSCRTFGCDGRMDSQVWDRCCVCGGDNSTCS 579
DB 109 -RVD---DGTGQCAATSR-----AVCSKGSCQIVGCDGLISSFRFDACGVCGRGDTCD 159
QY 580 PRKGSFTAGRAREYVTFLTVPNLTSTVIYANHRLFTHLAVRIGRYVYVAGKMSIPNTT 639
DB 160 --NGKF-----IW-----KVSEY----- 171
QY 640 YPSLLEDGVEYRVALTEDELPLESIRIWGPLOEDADIQVYRRYGBEYGNLTRDITFT 699
DB 164 -----TACASNCDDIVDSGAGRSIASTSQPIVVCVNAITGRVVPEKLC 215
QY 700 YFQPKRQAWAAVRGPCSVSC-----GAGLRWYNYS-----CLDQARKELVETQC 747
DB 172 -----TACASNCDDIVDSGAGRSIASTSQPIVVCVNAITGRVVPEKLC 215
QY 748 QGSQQPAPWPAEACVLEPCPPYMAVDFGPCSASCGGLRERPRVCVEAQQSLLKTLPPAR 807
DB 216 ADKLRPKVEARCPMLICPSRWAAADWTECVPHCGEGTRKEEVYCVQTAHNVVHVPTTF 275
QY 808 CRAGAQQPAVALETNCPQCPARWEVSEPPSCTSAGAGLALENETCVPGADGLEAPYTE 867
DB 276 CENGT-RPA-ABENCVSTSC-GRWEAGKWSKCTASCGQVRRRRHVACVCGGSDCDEG--- 328
QY 868 GPGSVDEKLPAPPEPCVGMSCPPQW---GHLDATSAGE-KAPSPWGSIRTGAQAAHWTPA 923
DB 329 GRPROETTCYAGIPCISATNSLDWDRAYLDGNTFGSMNDHNDWQAPRL---VAGEWS--- 383
QY 924 AGSCSVSGRGLMELRFLQW-----DSALRVVQE-----ELCGLASKPGSRRE 967
DB 384 --TCSSTCGTGYMSRTVECVAVNPISAPIKLPMSECCDQPKLFESCEVRSCLQEDS 441
QY 968 VQAVPCPARWOY-KLAACSVSGRGVVRRIYCARAHGEDDGBEILLDTCCOGLPR-PE 1025
DB 442 KLSDEAPFYQRYGDWQTQASCLGGKQKALKCIQV---STGKSQVW-SQCDARRRPPE 497
QY 1026 PQEACSLGPCPPRWKVMSLGPCASCGGLTARRSVACVQLDQGDQDEVEDEAACAAIV--- 1082
DB 498 KSRPCNQHCPPFWLTSTKYSDCSMSCSGTARRSVKC-----AQTVSKTDGDAHIVLRD 552
QY 1083 -----RPEASVPECLTADCTYRW-----HVGTMWECSSVCGDG:QRRRD 1120
DB 553 DRCHFKKPKQETECNVVACPAWTVWSLNGKHNKIKNLKNKTAQWTECSRDCSGERRRQV 612
QY 1121 TCLGPOQAQAPV---PADFOHLKPKVTVRGCWAGPC 1153
DB 613 WCEIRDSRGKTORRPDVECDANTKPTQVEVCSFGSC 648

```

RESULT 10
T00260      hypotheical protein KIAA0605 - human
C/Species: Homo sapiens (man)
C/Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C/Accession: T00260
R:Ngase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.
DNA Res. 5, 31-39, 1998

```

A,Title: Prediction of the coding sequences of unidentified human genes. IX. The complete coding sequence of the gene for human thrombospondin type 1 repeat homology
A,Reference number: Z14086; MUID:98290545; PMID:9628581
A,Accession: 700260
A,Status: preliminary; translated from GB/EMBL/DBJ
A,Molecule type: mRNA
A,Residues: 1-951 <NAG>
A,Cross-references: EMBL:AB011177; NID:g3043733; PIDN:BAA25531.1; PID:g3043734
A,Experimental source: brain
C,Genetics:
A,Note: KIAA0605
C,Superfamily: thrombospondin type 1 repeat homology
F,45-106/Domain: thrombospondin type 1 repeat homology <THR1>

	Query Match	7.1%; Score 546.5; DB 2; Length 951;
	Best Local Similarity	22.2%; Pred. No. 7.8e-27;
	Matches 216; Conservative 105; Mismatches 381; Indels 273; Gaps 39;	
QY	408 GRWSSWGPRSCRSRSGGGVTRRRRCNNRP----	PAFGGRACVCADLQAEMCNTQACEK 463
DB	51 GEWTN-----TAFSRGCGGVTSQERHCLQRKSVPGFNGRTCTGTSKRYQLCRVDECPP 107	
QY	464 TQLEFMSQQCARTDQQLRSSPGSAGFYHWGAAP----	HSQGDALCRMCAIGESFIM 519
DB	108 DGRSFREQC-----SPFNHYNGRTHQWKPLYDPDVVHISSEKPCDLHCTTVDGORQLM 162	
QY	520 KRGDSFLDTRCMPGSPREDGTLUSLVSGSCRFTGCDGRMDSQQVMDRCQVCGDNSTCS 579	
DB	163 VPA---RDGTSCKLTDLR-----GVCVSGKEPIGCDGLFSHTLDKCGICQOGDSSCT 214	
QY	580 PRKGSFTAGRAREYVTFLTVP-NLTSVYTANHRPLTHLAVR-IIGRVVVGAKMISPN 637	
DB	215 HVTGNRKGNAGHLGXSLVTHIPAGARDIQIVERKSADVLADEAGYFFNGNYKV--- 271	
QY	638 TTYPSELLDGRVEYEVALTEDRLP-----LEEIRMGPLQEDADIQTVVRRYGEEYG 689	
DB	272 -----DSPKNFIAGTVVKXRPMDVYETGIEYIVAQGFNOGLNVVMWNQNS-- 321	
QY	690 NLTRPDIFTY--POP----KPRQAWMAAVRGFCVS CGAGL----- 726	
DB	322 ---PSIIFEYLLQPHESSRPQIVYGFSESABESQGLDGAGLMGFPHNGSLYGQASSE 377	
QY	727 -----RWNTVSCLD-----QARKELVETV---OCQGSQQPPAWPEACVL----- 762	
DB	378 RLGLDNRLFHGHLDMELGPSQGETNEVCEQAGGACEGPPRGKGRDRNVTGTPLTGD 437	
QY	763 ---EPCCPYWAVGDGFGCSA-----SCGGGLRE---RPVRCVBAQGSLTKLP----- 804	
DB	438 KODBEVDTHFAECFPFSANAISDLLGAGSDLKDFTLNETHNSIFAQAPRSLARSFFV 497	
QY	805 -----PARCRAGAQQ 814	
DB	498 DYENEGAGPYLLNGSYLELSDRVANSSEAFPENVSTSLTTSAGNRTHKARTPKARK 557	
QY	815 PVALET CNPOCPA---RWEVSEPSCTSAGGAGLALENETCVPGADGLEAPVTEGP GS 871	
DB	558 QGVS-----PADMYRWKLSSHPECSTACTTGWSAYAMCV-RYDGEV-----DDS 602	
QY	872 VDEKLPAPBP-----CVGMSCPPGWCHLD-----ATSAGE-----KAPSP-----W 907	
DB	603 YCDALTREPPEVHFCAAGRCQRPWETSWSSECSTCGEGYQFRVVRCKWMLSPGFOSSVY 662	
QY	908 GSIRTGAQAAH-----VMTPAAG-----SCSVSCGRGLMEIRFLCMDSALRPVQ 952	
DB	663 SOLCEAAAVRPERKTCRNPCACPQWEMSEWSECTAKCGERSVVTR-----DIRCED 716	
QY	953 BELCGLAKPGSRREVCAVPCCPARWOYK-LAACSVSCGRGVVVRILYCARAHGEDDGE 1011	
DB	717 EKLCDNTPRPVGEKN-CTGPPCDRQWTVSDWGPCSGSCGQGRTHRVYCKTSDG-----R 770	
QY	1012 ILDTQCQGLPRPEOEAASLEPCPPRWKVMSLGPCSASGLGTARRSVACVOLDDQGDV 1071	
DB	771 VPVESOCOMETKPLATHPCCDKNCNCPAHLAADWERCNITCGRGVKRLLVLCMELANGKP 830	

584 SFTAGRAYEYTFLLTVTNLTISVYIANHRPLFTHLAVRIGGRYVAGKMSISPTNTPSL 643
 : : : : :
 594 -----PITQSSSVYRGQ----- 606
 644 LEDGRVEYRVALTEDRLRLEIRITWGLPQEDADIQVYRRYGEYGNLTPDITFTYFQP 703
 : : : : :
 607 -----WG----- 608
 704 KPROQAW-WAAAVRGPCSVSCAGLEWVNYSLDQARKELVETVQCGSQPPAWPEACVL 762
 : : : : :
 609 ---TWLSLWTS---CTATCGGKYKRNAC-----SITQCEGNEDE--TEVCSS 649
 : : : : :
 763 EFCPPYWAVG-----DFGPCSACGGGLRRPRVRCVEAQGSLILKTLTP----- 804
 : : : : :
 650 ESCPVLRVGNENSWTWTWNHCSVSCGRSQARYKCLSPHRTLAPDCPGENKVTNLR 709
 : : : : :
 805 -----PARC---RAGAQQPAVALETNCPQCPA-----RWEVSEPSSCTS 841
 : : : : :
 710 TFFKARSVMCSVRCKTKRNTISEKNIEVRSCDNGPCNAIGVWTGWGW-----STCST 764
 : : : : :
 842 AGGAGLALENETCVPGADGLEAPVTEGGSVDKLPAPPCVGV-----MSCPFGWGHLD 896
 : : : : :
 765 SCGPGLVVRQTC-----NREPCDGAHERSSCNV-----A 795
 : : : : :
 897 TSAGKASPPWGSIRTGAAAHVWTPAGSCSVSCRGMLMEIRFLICMDSALRPVQBELC 956
 : : : : :
 796 TCQNDGSIWNE-----WS-----DCSRVCGKGLRSRSRSCFGSG-----C 832
 : : : : :
 957 GLASRPGSRREVCOAVPCPAR-----WQYKLAACSVSCRGVVRILYCARAHGE 1006
 : : : : :
 833 MGAS---SEQQFNEQACASSANDWGTWSGW---SQCSVSCGAGVKRTRTCRTGN-- 883
 : : : : :
 1007 DDGEHLLDTQCGLPREPQPCACSLPCPPRWKVMSLGPCSASCGLGTARASVACVQLD 1066
 : : : : :
 884 -----CP-----GNVKSALC----- 894
 : : : : :
 1067 QCGQVEYDEAACAALVRPEASVPCLIADCTYRHHVGTWMECSVSCGDGIQRRADTCIG 1124
 : : : : :
 895 NDRDCENKNAWGG-----W-GYWSSCSETCGDGVRKVRKCVG 932
 : : : : :
 RESULT 12
 T16892
 hypothetical protein t19d2.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
 C:Accession: T16892
 R:Riabtsev, D.
 submitted to the EMBL Data Library, December 1995
 A:Description: The sequence of C. elegans cosmid T19D2.
 A:Reference number: Z16599
 A:Accession: T16892
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-860 <BEN>
 A:Cross-references: EMBL:U42846; NID:g1125809; PID:g1125810; PIDN:AAA83600.1; C:
 C:Genetics:
 A:Gene: CESP.T19D2.1
 A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2;
 Query Match 5.5%; Score 428.5; DB 2; Length 860;
 Best Local Similarity 19.8%; Pred. No. 2.1e-19;
 Matches 211; Conservative 109; Mismatches 321; Indels 425; Gaps 47;
 90 RQRQRRAAGGILHLELVAVGPDVF-----QAHQEDTRYVLTNINIGAEILLRDP 141
 : : : : :
 123 RRAEPHRR--DILATVELAVFADDAMWDHFKMYGKAENMHTFMVAVNNIDVLYTQR 179
 : : : : :
 142 SLGAGFRVHLVKVILTEPEGAPNI-----TANLTSLLSVCGHSQTINPEDDTPGH 194
 : : : : :
 180 LLQPRINIKIVRYEIL--KNIPHLNARKHNGVDRLDLDAFCQYQNEINPNPDADPRH 236
 : : : : :
 195 ADLVAVITRFDLELPDGNFQVRGTVTLGGACSPFTSLITEDGFQLVGTIAEIHGSFG 254
 : : : : :
 254 -----PITQSSSVYRGQ----- 266

Db 237 WDALLFSGYDLH-RNGKTVAGVAPVKMGSGVRSCTINEGLDRGVSFVVTHEMGHSLG 295
QY 255 LEHDG-----APGSGCGPSGHVWASDGAAPRAGLAWSPCSRRLQLLSL-----S 298
Db 296 MYHDGNECDLRCIMSPSVSGKT-----HWSQCSVNEMAFVGHGLGD 340
QY 299 AGRARCVWDPRPQPGSAGHPDQPGLYYANECRVAFGPKAVACTIFAREHLD----- 353
Db 341 FRPNCLQDASANEQRMVAFKESBPFGQLFTLDEQCEIFHGE-----CWKHEKDGQTMQ 395
QY 354 -MQQALSCHTDPLDQSSCSLLVLLDGTGCGVEKWCCKGRCSLVELTPIAAVHGRWSS 412
Db 396 NIQMWGCG-----NGEIVIRTAHALEGTYCGFMICRQCGVSSQLMRVTV---GGWST 449
QY 413 WGRSPCSRSCGG-----GVTRRRQCNPRPAFGGRACVAGADLQAEQMCNTQAC 461
Db 450 WNDRP--APTCCGRFCSOCEIRGQIRIMESIRQCNPNSSNNGGAPCGDEARGWVCHRDVC 507
QY 462 EKTQLE-FMQSQCAQTGQPIRSPFGGASFYHMGAAVPHS---QG-----DALCRMCA 512
Db 508 NGDSIENYATRVCSR-----LRDE-----NAIPNTILSGEGMQPEQAMCKIWCLI 552
QY 513 IGESFIMKRGDSFLDGTCTRCMPSPGREDGTLSLCVSGSCRTFTCGDGRMDSQQVWDRCOVCG 572
Db 553 SGTNI--RTVSNFPDGAFCGPG-----QYCIKGECPRL-----LCG 587
QY 573 GDNSTCSPRGSFAGRAREVVTFTVTPNLTSTVIANHRPLFTHLAVRIGGRVYVAGRW 632
Db 588 -----STTLAYSEADCPD----- 600
QY 633 SISFNNTYPSLLEDGRVEYRVALTEDRLPRLEIRINGFLOEDADIQVYRYGEEYGNLT 692
Db 601 SVLQTTITTP-----NPHVHSV-----DQFAGKT 624
QY 693 RPDITFTYFQPKPR---QAWVWAAVRGPGSCVSGAGLRWNVNSCLDQARKELVETVQCG 749
Db 625 NP-----YKEHKKTPFLNEMSGSWSEC-----VTYDCHTQGVK--VRVRRC-- 665
QY 750 SQOPPAWPEACVLEPCPYWAVGDPGCSASCGGGLRERPVRCVCAQGSLLKTLPPARCR 809
Db 666 -----LAGVCAGALRR----- 677
QY 810 AGAQPAVALETNCPQPCPARWEVSEPSSTCSAGGAGLALENETCPVGADGLEAPVTEGP 869
Db 678 -----QPC-----TRP--CT-----GSERPITTSF 695
QY 870 GSVDEK-----LPAPEPCVGMSCPPGWHILDATGAGEKAPSPWGSIRTCQAHAHVWTPA 923
Db 696 PQOTFRNRFTAPLNRQTNMILRKVDHWG-----PW----- 726
QY 924 AGSCSVSGRGLMBELFLCMLSALRVPVQBELCGLASKRPGSRREVC-----QAVPCP-- 975
Db 727 -SACSVTCGTG-QKLR-----RRENCIGQCAETGPCVMQ 759
QY 976 -----ARWQYKLAACSVSCGRGVV--RRILYCARAHGEDDGEIILLDTCCQGLPRP 1024
Db 760 SCRENKNTWTS--QWSDCVNCGEGVQFRKACFAAFCKGKSDSVRNCYGQRC----- 812
QY 1025 EPOEACSLPCPPR-----WKVMSLGPSCASGGLGTARRSVACVQ 1064
Db 813 --SETTTRPLVNRSLWTGHS--SWSSCSCTKCGIGQTRRRRCYQ 854

RESULT 13

T14764
Hypothetical protein DKFZp434H204.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14764
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18181

A:Accession: T14764
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-898 <NAM>
A:Cross-references: EMBL:AL110226
A:Experimental source: adult testis; clone DKFZp434H204
C:Genetics:
A:Note: DKFZp434H204.1
Query Match 4.6%; Score 354; DB 2; Length 898;
Best Local Similarity 20.0%; Pred. No. 1.2e-14;
Matches 169; Conservative 65; Mismatches 219; Indels 392; Gaps 36;
QY 703 PKPRQAWWAAVRGP-----CSVSCGAGLRWNVNSCLDQARKELVETVQCGSQOPPAWPEA 759
Db 31 PPPVFSWHY---GFWTKCTVTCGRGVQQRNVYCLER-QAGEVDEEHCDPLGRPDQQRK 85
QY 760 CVLEPCPYWAVGDPGCSASCG-GGLRERPVRCVCAQGS-SLLKTLPPARCRAGAQQPAV 817
Db 86 CSQPCPARWAGWQJLCSGCGPGGLSRRRAVLCSVGLDQSALEPPACE---HLPRP 142
QY 818 ALET-CNPQ-PCPARWEVSEPSSTCSAGGAG-----LALENETCPVGADGLEAPVTE--- 867
Db 143 PTEPCNRHVPCPATWAGNWSQSVTCGEGTQRENVLCTNDTGV-CDEAQOPASEVTC 201
QY 868 -----GPGSVDEKL-----PAPEPCVGMSCPPG----- 890
Db 202 SLPLCRWPLGTLPESGSGSSSHELFEADFIPIHLAPRPPSPASSPKPGTGMGNAIEEEA 261
QY 891 -----WG-----HLDATSAGEK----- 902
Db 262 PELDLPGEVFDYFYDYNFINTHEDLSYGFSEEPDLDTAGTDRTPPPHSRPAAPSTGS 321
QY 903 -----APSPWGS----- 909
Db 322 PVPATEPPAAKEGVLFGWSPSPWPSQAQGRPPPPSEQTGPNPLINFLBEDTPIGAPDL 381
QY 910 -----IRTCG----- 914
Db 382 GLPSLSWRVSTDLQTPATPESQNDFFVQKDSQSLQPPPPWRDRTNVEVFKDDEPKGRGA 441
QY 915 -----QAHHVWT-----PAAGS 926
Db 442 PHLPPRPSTLPLSPVSGSTHSSPSPDVAELWTGTVAVPALEGGGLGVDSSELWPTVG 501
QY 927 CSV-----SCRGLMELRFLCM-----D 944
Db 502 ASLLPPPIAPILPEMKVRDSSLLEPGTPFPAPGPGSWDLQTVAVMGTFLLPTTLTGLHMP 561
QY 945 SAL-----RVPQBELCGLAS--KPGSRREVCQAVPCP----- 975
Db 562 PALNPGKQGPESLSPEVPLSSRLSTPAWDSFANSHRVFETQPLAPSLAEAGPPADPLV 621
QY 976 ---ARWQY-KLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTCCQGLPRPEBSACS 1031
Db 522 VRNASWQAGNWESECTTCGLGAVWEPVRCSSGRDED-----CAPAGRQPARRCH 671
QY 1032 LEPCCPKNWMSLGFCSASCGLGTARRSVACVQLDQGGQDVE-VDEAAC-AALVVRPEASVP 1089
Db 672 LRPC-ATWHSNGWSKSRSCGSGSVRDVQCVDT---RDLRLPLRPHFCQPGPAKPPAHRP 727
QY 1090 CLIADCTYRHHVGTWMECSVSCGGIGIQRRTDTCGLGPAQAPVADFCQHLPKVTVVREGCW 1149
Db 728 CGAQPC-LSWYTSRWRCEACGGEQQLVTC-----PEFG-LCEEALRNTTTRPCN 778
QY 1150 AGPCV-----GO-----GTPSLVPHHEAAAAGRTTATPAGA--CGR 1183
Db 779 THPTCTQVVGWPGQCSAPCGGVQRLVKVNTQTGLPEEDSDQCGH-EAWPSSRRPCGT 837
QY 1184 QHLEP 1188
Db 838 EDCFP 842

RESULT 14

T25061
hypothetical protein T21B6.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C/Accession: T25061
R/Cottage, A.
submitted to the EMBL Data Library, November 1995
A/Reference number: Z19975
A/Accession: T25061
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-788 <WIL>
A/Cross-references: EMBL:Z68011; PIDN:CAA92014.1; GSPDB:GNC00028; CESP:T21B6.3
A/Experimental source: clone T21B6
C/Genetics:
A/Gene: CESP:T21B6.3
A/Map position: X
A/Introns: 20/1; 47/1; 76/1; 152/1; 735/2; 754/2

Query Match 3.5%; Score 275; DB 2; Length 788;
Best Local Similarity 23.1%; Pred. No. 1.1e-09;
Matches 112; Conservative 40; Mismatches 149; Indels 184; Gaps 28;

QY 718 CSVSCGAGLRWVNSCLDQARKELVETVQCQSQPPAPWPEACVLEPCPPYMAVDFGFC 777
Db CSCTCGDGAK-----SRRRCSTNNCGADYE-----TEPCNLGPGQWSEWSTC 514

QY 778 SASCGGGLRRPVRVCEAQAQSLKTLPPARAGAAQPAVALETCPNPPCP--ARWEYSE 835
Db SASCGSQRETRFC-----HLGNRCCKDYES-----EQCSAGPCPWSQWE-D 559

QY 836 PSCTAGAGLALLENCTVPGAGD---LEAPVTEGPGSVDEKLPAPPCVGMSCPPGHWG 892
Db WGCSTVTCGGGAVRQRTCLGGVFGDHLCCQPKTE-----QRACDGGPCL-W- 606

QY 893 HLDATSAKAPSPWGSVIRTAQAQAAHWTTPAGSCSVSCGRGLMELRLCMDALRVVQ 952
Db -----SPW-----QEWS-----TCSASCGSKMR-----R 626

QY 953 EELCGLASK---PGRREVQAVPCPARWQ--YKLAACSVSCGRGVWRRLYCARRAGHD 1007
Db QYVQCFQDTCQGNBESQFCVGPFC-AEWTEWCEWSSGSSCKCPQRTTRCLGPNQGE 685

QY 1008 DGEILLDTQCGLRPP-----EPQBACS-LEPCPPRWKVMYSLGPCSASGGLTARRS 1059
Db -----ATTQCG---PSIETTLCEGQSCCNWSEW--HWSM-----CDKECGGQVRY- 727

QY 1060 VACVQLDQGDVEVDERAACALVRPEASVPCLIADCTYRHWVG--TWMECS--VSCGDGIQ 1116
Db -----IYMFRTGCEWSPCSTQLACEVGQV 752

QY 1117 RRDTCILGPQAQAPVPADFCQHLPKPVTVRGCMAGPCVCGQTPSLVPHEEAAAPGRTTAT 1176
Db -----GC-----HCIG-----LAESQCRGLTQCP 783

QY 1177 PAGAC 1181
Db 784 PKPFC 788

RESULT 15

T00026
brain-specific angiogenesis inhibitor 1 - human
N/Alternate names: BAI1 protein
C/Species: Homo sapiens (man)
C/Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 12-Feb-1999
C/Accession: T00026
R/Nishimori, H.; Shiratsuchi, T.; Umano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida
submitted to the EMBL Data Library, June 1997
A/Reference number: Z14064
A/Accession: T00026

A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1584 <NIS>
A/Cross-references: EMBL:AB005297; NID:dl1175078; PID:dl024528
A/Experimental source: brain
C/Genetics:
A/Gene: GDB:BAI1
A/Cross-references: GDB:9838088; OMIM:602682
A/Map position: 8q24-8q24
C/Superfamily: thrombospondin type 1 repeat homology
F:408-462/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 3.4%; Score 264.5; DB 2; Length 1584;
Best Local Similarity 24.0%; Pred. No. 1.1e-08;
Matches 164; Conservative 61; Mismatches 242; Indels 217; Gaps 43;

QY 697 TPTYFQPKRQAVMAVRG---PCSVSCGAG-LRWVNY-SCLDQARKEL-VETVQ---- 746
Db SWTLRNPDPRTYLYMKVAKAPVPCS---GPRVRYTQFDFLESTRYLGVSEDEVLR 125

QY 747 -CQGS-----QOPP-----AWPEACVLEPCPPYMAVDFGFC 780
Db LCDPSAPLAPLQASKQFLQMRQPPQHDGLRPRAGPPGPTDDFSVEILVGNRPRAA 185

QY 781 C-----GG-----GLRERVRCV--EAQSLKLTLPAR--CRAGAQQPAVA 818
Db CQMLCRWLDAAGLRSSHPGIMQTPCACLGGGAGGPAAGPLAPRGDVC---LRDAVA 241

QY 819 LETCNTPQC-----PARWEV-SEPSCSTAGAGLALLENCTVPGADLEAP 864
Db ---GGPENCILSTLQDRGHGATGWLKLSLWGCTRCGGGLQTRTRCLP-----AP 292

QY 865 VTGPGSVDEKLPAPEPCVGMSCPPGWHLDATAGAKA-----PSP 906
Db GVEG-GGCEGVLEBEGROCNREACPA-GRSSRSQSLSRTDARRREELGDELQOQGFAP 350

QY 907 WGSIRTAQAQAAHWTTPAGSCSVSCGRGLMELRLCMDALRV-----PVQEE-LCGLASK 961
Db ---QIGDPAAEWSPWS-VCSSTCGEGWQTRTFVSSVSTQCSGPLEQRCLC----- 400

QY 962 PGRREVQAVPCPARWQ--YKLAACSVSCGRGVWRRLYCARRAGHDGEEILLDTQCQ 1019
Db ---NNSAVC---FVHGAWDEWSPWSLCSSTCGRFRDRTTRTCRPPQ-----FGNPCE 447

QY 1020 GLRPEPQ-EACSLPECPPR-----WKYMSLGPCSASGCGLTARRSVACVQLD--- 1066
Db G---PEXQTRFCNTALCFGRAVDGNWNEWS--SWSACSASCSQGRQQRTRCNGRPSYVGA 502

QY 1067 --QSQDVEVDEAACALVRPEASVPCLIADCTY--RHWV-GTWMECSVSCGDGIQRRRD 1121
Db ECQGHWETRD-----CFLOQCPVDGKQAWASGWSGCSVTGAGSQRRERV 548

QY 1122 CLGP---QAQAPVPAD-----FCQHLPKPVTVRGCMAGPCVCGQTPSLVPHEEAAAPGRT 1173
Db CSGLPFGGAACQGPQDEYRQCGTQRCPEPHEI-----CDBDNFGAVIWKETPAGEVA 601

QY 1174 TATPAGACG---RQHLEPTGTIDMRGPGQADCAVAI-----GR 1208
Db VRCPRNATGLILRRCELDEGIAYWEPTTYIRC-VSIDYRNIQMTTREHLAKAQRGLPGE 660

QY 1209 PLGEVW-TLRVLESSLNCAGDML 1231
Db GVSEVIQTLVEISQDGTYSYGDLL 684

Search completed: March 13, 2004, 07:44:34
Job time : 57.4125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 29.5 Seconds
(without alignments)
2499.367 Million cell updates/sec

Title: US-09-836-712-2

Perfect score: 7751

Sequence: 1 PGRPTRKAPSHSAPLLGLA.....LQSWPEMQDPQSWKKEGT 1416

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1224.5	15.8	1906	1 AT20_MOUSE	P59511 mus musculus
2	1189	15.3	1911	1 AT20_HUMAN	P59510 homo sapien
3	1167.5	15.1	1935	1 AT20_HUMAN	Q9p2n4 homo sapien
4	1146.5	14.8	1224	1 AT16_HUMAN	Q8te57 homo sapien
5	1126	14.5	1593	1 AT12_HUMAN	P58397 homo sapien
6	1040.5	13.4	1081	1 AT18_HUMAN	Q8te60 homo sapien
7	1026	13.2	1077	1 AT17_HUMAN	Q9h324 homo sapien
8	1005	13.0	1095	1 AT17_HUMAN	Q8te56 homo sapien
9	978	12.6	1213	1 AT22_MOUSE	Q8c9w3 m adamus-2
10	956.5	12.3	1223	1 AT14_HUMAN	Q8wx88 homo sapien
11	955	12.3	1205	1 AT22_BOVIN	P79331 b adamus-2
12	945.5	12.2	967	1 AT21_RAT	Q9wuq1 rattus norv
13	944.5	12.2	968	1 AT21_MOUSE	P97857 mus musculus
14	941	12.1	1211	1 AT22_HUMAN	Q95450 h adamus-2
15	914	11.8	967	1 AT21_HUMAN	Q9uh18 homo sapien
16	907	11.7	1205	1 AT23_HUMAN	O15072 homo sapien
17	893.5	11.5	950	1 AT15_HUMAN	Q8te58 homo sapien
18	867.5	11.2	997	1 AT27_HUMAN	Q9ukp4 homo sapien
19	852	11.0	1210	1 AT19_MOUSE	P59509 mus musculus
20	847	10.9	1207	1 AT19_HUMAN	Q8te59 homo sapien
21	830	10.7	890	1 AT28_HUMAN	Q9up79 homo sapien
22	768.5	9.9	930	1 AT25_HUMAN	Q9una0 homo sapien
23	761	9.8	905	1 AT28_MOUSE	P57110 mus musculus
24	758.5	9.8	930	1 AT25_MOUSE	Q9r001 mus musculus
25	746.5	9.6	837	1 AT24_HUMAN	O75173 homo sapien
26	721	9.3	630	1 AT24_RAT	Q9esp7 rattus norv
27	655.5	8.5	562	1 AT15_MOUSE	P59384 mus musculus
28	626	8.1	860	1 AT26_HUMAN	Q9ukp5 homo sapien
29	489	6.3	525	1 AT11_HUMAN	Q8n566 homo sapien
30	427.5	5.5	450	1 AT10_MOUSE	P58459 mus musculus
31	377	4.9	377	1 AT11_MOUSE	Q8b110 mus musculus
32	354.5	4.6	207	1 AT25_BOVIN	Q9tt92 bos taurus
33	304.5	3.9	867	1 SSGPO_BOVIN	P98167 bos taurus

ALIGNMENTS

RESULT 1

ID	AT20_MOUSE	STANDARD;	PRT;	1906 AA.
AC	P59511;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-TS20).			
DE	TS20.			
GN	ADAMTS20.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=22566039; PubMed=12562771;			
RA	Llamazares M., Cal S., Quesada V., Lopez-Otin C.;			
RT	"Identification and characterization of ADAMTS-20 defines a novel subfamily of metalloproteinases-disintegrins with multiple thrombospondin-1 repeats and a unique GON domain.";			
RT	J. Biol. Chem. 278:13382-13389(2003).			
RL	[2]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.			
RN	STRAIN=DBA/2;			
RC	MEDLINE=22806432; PubMed=12925592;			
RX	Rao C., Foerzler D., Loftus S.K., Liu S., McPherson J.D.,			
RA	Jungers K.A., Apte S.S., Pavan W.J., Belter D.R.;			
RT	"A defect in a novel ADAMTS family member is the cause of the belted white-spotting mutation.";			
RT	Development 130:4665-4672(2003).			
RL	Development 130:4665-4672(2003).			
CC	- FUNCTION: May play a role in tissue-remodeling process occurring in both normal and pathological conditions.			
CC	- COFACTOR: Binds 1 zinc ion per subunit (By similarity).			
CC	- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).			
CC	- ALTERNATIVE PRODUCTS:			
CC	Event-Alternative splicing; Named isoforms=2;			
CC	Name=1; Synonyms=ADAMTS20 B long isoform;			
CC	Isoid=p59511-1; Sequence=Displayed;			
CC	Name=2; Synonyms=ADAMTS20 A short isoform;			
CC	Isoid=p59511-2; Sequence=VSP_007606, VSP_007607;			
CC	- TISSUE SPECIFICITY: Expressed at low level in testis and brain.			
CC	- PFM: The precursor is cleaved by a furin endopeptidase (By similarity).			
CC	- DISEASE: Defects in ADAMTS20 are the cause of the belted (bt) phenotype. It is a pigmental defect which occurs as a result of a defect in melanocyte development.			
CC	- SIMILARITY: Belongs to peptidase family M12B.			
CC	- SIMILARITY: Contains 1 disintegrin-like domain.			
CC	- SIMILARITY: Contains 1 GON domain.			
CC	- SIMILARITY: Contains 15 TSP type-1 domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			


```

FT FT VARSPLIC 1505 1911 isoform 2).
FT FT /FTid=VSP_007107.
FT FT Missing (in isoform 2).
FT FT /FTid=VSP_007108.
FT FT T -> V (IN REF. 2).
FT FT E -> Y (IN REF. 2).
FT FT T -> M (IN REF. 2).
FT FT D -> E (IN REF. 2).
FT FT I -> V (IN REF. 2).
FT FT MISSING (IN REF. 2).
FT FT LILQ -> ILIE (IN REF. 2).
FT FT GMLLAK -> WHVIG (IN REF. 2).
FT FT S -> Q (IN REF. 2).
FT FT T -> R (IN REF. 2).
FT FT T -> A (IN REF. 2; CAD56159).
FT FT E -> O (IN REF. 2; CAD56159).
SQ SEQUENCE 1911 AA; 214656 MW; CF592E220D32B250 CRC64;

Query Match
Best Local Similarity 15.3%; Score 1189; DB 1; Length 1911;
Matches 339; Conservative 150; Mismatches 511; Indels 218; Gaps 40;

QY 67 QAVSYLSPGAPLKGRPPSPGFORQORRAAGGIIHLLELLVAVGPDVFOAHQEDTTRY 126
DB 230 ERLVGHSTKNVPLKD-----ERRHSRKKRLISYPRYIEIMVTADAKVVSAGHNLQNY 282
QY 127 VLTNLNIGALLRDRSLGAFRVLKVMVILTEPEGAPNITANLTSSLLSVCGWSQTINP 186
DB 283 ILTMSIVATYKDPISGNIHIVVVKLVIMTHREEEGFVNFDFGATLKNFCSMQQTQND 342
QY 187 EDDTDGADLVLYITRFDLELPGNRQVRGVTOLGGACSPWLSCLITEDTGFDLGVITIA 246
DB 343 LDDVHSHDHTAVLTITREDTCSSEKENTLGLSLVLTICDPLQSCFINEEKGLSAFTIA 402
QY 247 HEIGHSGLEHDGAPSGCG-----PSGHVMASDGAAPRAGLAWSPCSRRLSLLSAGRA 302
DB 403 HELGHTLGVQHDNDP--RCKEMKTKYKVMAPALSFHSPWSWNSCRKYVTEFLDTGVG 460
QY 303 RCWDDPPRPOFGAGHPDPAQGLYYSANECRVAFGPKAVACTFARHLDWCQALSCHT 362
DB 461 ECLLDKPDDE--IYNLSELPISGRYDGNKOCELAFGFGSQC-----PHNICMHLWCTS 513
QY 363 DPLDQSSGRLVPLLDGTGCGVEKWSKGRCSLIVELTPIAAVHGRWSSGPRSPCGRS 422
DB 514 TEKLHGKCFTHQVPPADGTGCGPMCHRHGJC--VNKETETRPVNGEWGPEYSSCRT 571
QY 423 CGGVTTRRQCNPRPAFGACRACVADLOAEMCNTOACEKTQLEFMGQOQARTDGOPLR 482
DB 572 CGGIESATRCRNPENRNGNYCVRRMKFRSCNTDSCPKGTQDFREKQCSDFNGKHL 631
QY 483 SPPGGASFYHNGAAVPHSQGDAL---CRHMCRAIGESFIMKRGDSFLDGTROMPSGPRED 539
DB 632 IS-GIPSNVRW---LPRYSIGTKDRCKLYQVAGTNYFYLLKDMVEDGTPCGTE---- 682
QY 540 GTLSLCVSGSCRTFGCGRMDSQQWDBCVQCGDGNSTCSPRKGSFTAGRAEYVTLTV 599
DB 683 -THDICVQCGMAAGCDHVLNSSAKIDKCGVCGDGNSSCKITGVFNESH-YGVNVVKI 740
QY 600 TPNLTSV---YIANHRLPFLHVR-IGGRVYVAGKMSISPNITYPSLLEDGR---VEY 651
DB 741 PAGATNVDIRQYSYSGQFDDSLSDAEGNLFNGNELL--TSKKEINVQGTTVLEY 798
QY 652 -----RVALT--EDRLPBLEIRIWG-----PLQBDADIQVYRYGEEYG 689
DB 799 SGNNAVERINSTROEKELLQVLCVGNLYNPVDVHYSFNIFLEERSDMFTWDPYGPWEG 858
QY 690 -----NLTPDITFTY-----FQPKP-----RQAWVAAVVRGPCS 719
DB 859 CTRMCQGLQERNITCIHKSDBSVSDKECDLPLPSFVTQSCNTPCRLRGMLLAKKSECS 918
QY 720 VSCGAGLRWNYSLDQAKELVETVQCO----GSOQPPAWPEACVLPFCPPYVAVGDFG 775
DB 919 SQCGQGYRTLDIHCWKYSIHE-GQTVQVDDHYCYGQDLKPPTQELCHGNCVCFTRWHYSEWS 977

```

RESULT 3

ATS9 HUMAN

```

ID ATS9 HUMAN STANDARD; PRT; 1935 AA.
AC Q9PNA; Q9NR29;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).
GN ADAMTS9 OR KIA1312.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Fetal;
RX MEDLINE=20396138; PubMed=10936055;
RA Clark M.E., Keiner G.S., Turbeville L.A., Boyer A., Arden K.A.,
RA Maki R.A.;
RT "ADAMTS 9, a novel member of the ADAM-TS/Metallospodin gene
RT family.";
RL Genomics 67:343-350(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wight T.N., Leduc R., Apte S.S.;
RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RT subfamily related to Caenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513(2003).
RN [3]
RP SEQUENCE OF 159-1935 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT the complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";

```

RNA Res. 7:65-73 (2000).

-!- FUNCTION: Cleaves the large aggregating proteoglycans, aggrecan and versican.

-!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1838-Glu-|-Ala-1839 site and versican at the 1428-Glu-|-Ala-1429 site.

-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

-!- ALTERNATIVE PRODUCTS:

Event=1; Synonyms=ADAMTS-9B;

Isoid=Q9P2N4-3; Sequence=Displayed;

Name=2; Synonyms=Long;

Isoid=Q9P2N4-1; Sequence=VSP_007548, VSP_007549;

Note=May result from the retention of an intron in the cDNA leading to a premature stop codon;

Name=3; Synonyms=Short;

Isoid=Q9P2N4-2; Sequence=VSP_005499, VSP_005500;

-!- TISSUE SPECIFICITY: Highly expressed in all fetal tissues. Expressed in a number of adult tissues with highest expression in heart, placenta and skeletal muscle.

-!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix (By similarity).

-!- DOMAIN: The ancillary domains, including the TSRs domain, are required for specific extracellular localization and for its versicanase and aggrecanase activities.

-!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

-!- SIMILARITY: Belongs to peptidase family M12B.

-!- SIMILARITY: Contains 1 disintegrin-like domain.

-!- SIMILARITY: Contains 1 GON domain.

-!- SIMILARITY: Contains 15 TSP type-1 domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL; AF261918; AAF89106.1; -.

EMBL; AF488803; AAO15765.1; -.

EMBL; AB037733; BAA92550.1; -.

HSP; F15167; IATL.

MEROPS; M12.021; -.

Genew; HGNC:13202; ADAMTS9.

MIM; 605421; -.

GO; GO:0008237; F:metallopeptidase activity; TAS.

GO; GO:0007275; P:development; TAS.

GO; GO:0006516; P:glycoprotein catabolism; TAS.

InterPro; IPR001762; Disintegrin.

InterPro; IPR001818; Pept_M10A_M12B.

InterPro; IPR006025; Pept_M_Zn_BS.

InterPro; IPR001590; Peptidase_M12B.

InterPro; IPR002870; Peptidase_M12B_N.

InterPro; IPR000884; TSPI.

Pfam; PF01562; Pep_M12B_propep; 1.

Pfam; PF01421; Repolysin; 1.

Pfam; PF00090; tsp_1; 11.

SMART; SM00209; TSF1; 12.

PROSITE; PS00215; ADAM_MEPPO; 1.

PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.

PROSITE; PS00437; DISINTEGRIN_1; FALSE_NEG.

PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.

PROSITE; PS00092; TSPI; 14.

PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydroxylase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Extracellular matrix; Alternative splicing.

FT SIGNAL 1 18 POTENTIAL.

FT PROPEP 19 287

FT CHAIN 288 1935 ADAMTS-9.

FT	DOMAIN	293	499	METALLOPROTEASE.
FT	DOMAIN	509	587	DISINTEGRIN-LIKE.
FT	DOMAIN	588	643	TSP TYPE-1 1.
FT	DOMAIN	644	752	CYS-RICH.
FT	DOMAIN	753	877	SPACER.
FT	DOMAIN	878	936	TSP TYPE-1 2.
FT	DOMAIN	939	997	TSP TYPE-1 3.
FT	DOMAIN	998	1049	TSP TYPE-1 4.
FT	DOMAIN	1052	1109	TSP TYPE-1 5.
FT	DOMAIN	1110	1166	TSP TYPE-1 6.
FT	DOMAIN	1182	1240	TSP TYPE-1 7.
FT	DOMAIN	1241	1296	TSP TYPE-1 8.
FT	DOMAIN	1328	1379	TSP TYPE-1 9.
FT	DOMAIN	1382	1440	TSP TYPE-1 10.
FT	DOMAIN	1441	1494	TSP TYPE-1 11.
FT	DOMAIN	1497	1555	TSP TYPE-1 12.
FT	DOMAIN	1556	1611	TSP TYPE-1 13.
FT	DOMAIN	1612	1676	TSP TYPE-1 14.
FT	DOMAIN	1677	1734	TSP TYPE-1 15.
FT	DOMAIN	1735	1935	GON.
FT	DOMAIN	88	96	POLY-SER.
FT	SITE	223	223	CYSTEINE SWITCH (POTENTIAL).
FT	METAL	434	434	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT_SITE	435	435	BY SIMILARITY.
FT	METAL	438	438	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	444	444	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	135	135	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	271	271	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	749	749	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	840	840	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1213	1213	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1267	1267	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1788	1788	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1806	1806	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLIC	1064	1072	CLVTGCGKH -> VRWEGCYFP (in isoform 3).
FT	VARSPPLIC	1073	1935	/FtId=VSP_005499.
FT	VARSPPLIC	1624	1629	Missing (in isoform 3).
FT	VARSPPLIC	1630	1935	/FtId=VSP_005500.
FT	VARSPPLIC	1630	1935	CSVTCG -> VPSWEL (in isoform 2).
FT	VARSPPLIC	1630	1935	Missing (in isoform 2).
FT	CONFLICT	46	46	S -> G (IN REF. 1).
FT	CONFLICT	96	96	P -> S (IN REF. 1).
FT	CONFLICT	182	182	D -> G (IN REF. 2).
FT	CONFLICT	367	367	F -> L (IN REF. 1).
FT	CONFLICT	1117	1117	V -> G (IN REF. 3).
SQ	SEQUENCE	1935	AA; 216556	MM; FD3D51E88300A3C6 CRC64;

Query Match 15.1%; Score 1167.5; DB 1; Length 1935;

Best Local Similarity 25.4%; Pred. No. 9.3e-64;

Matches 347; Conservative 170; Mismatches 506; Indels 343; Gaps 50;

QY	89	QRQQRQRAAGGILHLELLAVAGPDVFQHQEDTERVLTNLNIGALLRDPISLGAQRF	148
DB	279	KTHRTKRPFLSYPRFVFLVADNRNVMVHGENLOHYLLTMSIVASYKDPISGNLN	338
QY	149	VHLVKWILTEPEGAPNITANLTSLLSVGCWSOTINPEDDTPG--HADLVLYITRFDL	206
DB	339	IVIVNLIVIHNEODGPSISFNACTTLKNFCQWQHSKN-----SPGIHHDVAVLTRODI	393
QY	207	ELPDGHRQVRGVTQLGACSPFTWSCLTETDPTGLGVYTHIAEIGHSGLEHDPGSGCG	266
DB	394	CRAHDKCDTLGLAELGICDPYRSCSISDSGLSTAFTHAELGHVFNPHD--DNKKCK	451
QY	267	PSG-----HYMASDGRAPRAGLAWSPCSRQLLSLSAGRCARCVDPFRPQSGAGHP-P	320
DB	452	EEGVKSPQHYMAPTLNFYTNPMWMSKSRKYITFTDGTGYCECLNBPESR-----YPLP	507
QY	321	DAQPGLYSANEOCRVAFGPKAVACTFAREHLMCMQALSCHTDPLDQSSCSRLLPULLG	380
DB	508	VQLPGILYNNKOCLEIFGPGSVCPYNNQ-----CRLMNCNNVGHVKGRTOHTPWADG	563

Qy 381 TECGVEKMCCKGRSLVELTPIAAVGHWSNGPRSPCSRCGGGVVTRRRQNNRPPA 440
Db 564 TECEPGKCKYKFC--VPKEMDEVTDGSGWSPFGTCSRTCGGKTAIRECNPEPK 621
Qy 441 FGGRACVADLAEMCNTQACEKTQLEFMSQCCARTDQQLRSPGSGAFVHGAAPHSS 500
Db 622 NGGKYCVGRMRKFKSCNTEPCLQKQKDFRDEQCAHFDGKHFNIN--GLLPNVRW---VPKY 677
Qy 501 QGDAL---CRHMCRAIGESFIMKRGDSFLDGTCTCMPSPGPDGDTLSLCVSGSCRTFGCDG 557
Db 678 SGLMKDRCKLCFRVAGNTAYQLRBRVDGTPC-----GQDNDICVQGLCQAGCDH 731
Qy 558 RMDSQVMDRCVCGGDNSTCSRKSGFTAGRAREYVTLVTLNLTSLVVIANRPLFT- 616
Db 732 VLNSKARRDKCGVCGGDNSCKTVAGTFTNVH--YGYNTVVRKIPAGATNIDVRQHS--FSG 788
Qy 617 -----HLAVR-----IGRRVVA--GKMSISPTNTYPSLLEDCRVEYRVALTEDRLPR 662
Db 789 ETDDNNYALSSKGFLLNGFVVTMAKREIR-----IGNAVVEYSGSET-----A 835
Qy 663 LBEIRINGPLQEDADTQVRYRYGEEYGNLTPDITFY---FQPKPQAQWV-----711
Db 836 VERINSTRIEQLLELLQVL-----SVGLKYNPDVRYSNFIPIEDKPOQ--FYWNHSHGPWQA 889
Qy 712 -----AAVRGPCSV 720
Db 890 CSKPCQGERKRLVCTRESQDVLTSVDSORCDRLPOPGHITPCGTDCLDLRHWVSRSECSA 949
Qy 721 SCGAGLRWN-----YSCILDQARKELVETVCCGSGQPPAPWPEACVLEPCPPYVAWGDFG 775
Db 950 QCGGLGYRTLDIYCAKYSRLD--GKTEKVDGFC--SSHPKPSNRKCSGBCNCTGGWRYSAWT 1007
Qy 776 PCSASCGGLRBPVTCVEAQSLLKTLPRARCAGAQOPAVALETNCPQCPARWEYSE 835
Db 1008 ECKSCDGGTQRRRAICVNRNDV---LDDSKC---THQEKVTIQCSEFFPCP--QWKSGD 1060
Qy 836 PSSCTSAGAGLALNETCVPGADGLEAPVTEGPGSVDEKLPAPPCVGMGSCP-----PGW 891
Db 1061 WBECLVTCGKHGRQWQCFGEDRLNRMCD-----PEFKPTSMQTCQPECSAW 1111
Qy 892 GHLDATSAGEKAPSPGSIKTAQAARHWTTPAAGSCSVSCGRGLMELFLCWDLSALRPV 951
Db 1112 -----QAG-----PW-----VQCSVTCGGYCLRAVKCIIGYMSV 1143
Qy 952 QBELOGLASKPGRRREVQAVPC---PA-----RWQY--KLAACSVSCGRGV 993
Db 1144 DDNDNCNAATPTDQD--CELPSCHPPAPABETRRSTYSAPRTQWRFGSWTPCSATCGKGT 1202
Qy 994 VRILYCARAHGEDDGEIILLDTQCGLPRPEPOEACSLPCPRKWTMSLGPSCSASGL 1053
Db 1203 RMYVSC-----RDENGSADEACATLPRVAXECSVTFC--GQWKALDWSSCSVTCGQ 1256
Qy 1054 GTARRSVACVOLQGGQDVEDEAACAALVRPEASVPLIADCTVR-----1098
Db 1257 GRATQVMCVNY---SDHVIDRSECDQDIYPETQDCSMSPCQPTPSGLAQHFFQNE 1313
Qy 1099 -----MHVGTWMECSVSGDGIQRRRDTCGLGQAQAPVADPCQHL 1140
Db 1314 YRPSRASPSTHVLGNGQWRTGPMWAGCSSTCAGSQRRVVVC---QDENGVTANDCVERI 1370
Qy 1141 KPVTVRGWAGPCV-----GQGTFSL-----1161
Db 1371 KPDEQACSGPCQWAYGNWGECKLCCGGGIRTLVVCQSRNGERFPDLSCIELDKKPPD 1430
Qy 1162 -----VPHEAAAPGRITATPAGACR-----QHLE-----PTG 1190
Db 1431 REQCNTHACPHDAWSTGPMSSCSV--SCGRGHKORNVYCMADGSHLESYDCKHLAKPHG 1489
Qy 1191 TIDMRG-----PGQADCAVATGRPLGEVTLRLVLESSLNCASG 1228
Db 1490 HRKCRGRCPCFKWAGANSQCSVSGR-----GVQQRHVCCQIG 1527

RESULT 4
AT16 HUMAN
ID AT16 HUMAN STANDARD; PRT; 1224 AA.
AC Q8TE57; Q8TE57;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE ADAMTS-16 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 16) (ADAM-TS 16) (ADAM-TS16).
GN ADAMTS16 OR KIAA2029.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21356482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V., Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";
RT Gene 283:49-62(2002).
RL [2]
RN SEQUENCE OF 204-1224 FROM N.A. (ISOFORM 1).
RP TISSUE=Brain;
RC Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8TE57-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8TE57-2; Sequence=VSP_007664, VSP_007665;
CC -!- TISSUE SPECIFICITY: Expressed in fetal lung and kidney and in adult prostate and ovary.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix (By similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 6 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL: AJ315734; CAC86015.1; -.
CC EMBL: AB095949; BAC23125.1; -.
CC NEROPS; M12.026; -.
CC Genew; HGNC:17108; ADAMTS16.
CC MIM; 607510; -.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC Pfam; PF01562; Pep_M12B_prosep; 1.
CC Pfam; PF01421; Reprlysain; 1.

DR Pfam: PF00090; tsp_1; 5.
 DR PRINTS: PR01705; TSP1REPEAT.
 DR SMART: SM00209; TSP1; 6.
 DR PROSITE: PS00215; ADAM_MEPRO; 1.
 DR PROSITE: PS00446; CYSTEINE_SWITCH; FALSE_NEG.
 DR PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE: PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE: PS00900; PLAC; 1.
 DR PROSITE: PS00902; TSP1; 5.
 DR PROSITE: PS00142; ZINC_PROTEASE; FALSE_NEG.
 DR Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Alternative splicing.
 FT SIGNAL 1 24
 FT PROPEP 25 279
 FT CHAIN 280 1224
 FT DOMAIN 280 495
 FT DOMAIN 496 585
 FT DOMAIN 586 641
 FT DOMAIN 642 746
 FT DOMAIN 747 873
 FT DOMAIN 874 922
 FT DOMAIN 927 987
 FT DOMAIN 998 1048
 FT DOMAIN 1051 1115
 FT DOMAIN 1127 1181
 FT DOMAIN 1186 1223
 FT SITE 249 249
 FT METAL 433 433
 FT ACT SITE 434 434
 FT METAL 437 437
 FT METAL 443 443
 FT CARBOHYD 156 156
 FT CARBOHYD 310 310
 FT CARBOHYD 741 741
 FT CARBOHYD 780 780
 FT CARBOHYD 835 835
 FT CARBOHYD 905 905
 FT CARBOHYD 935 935
 FT VARSPLIC 1063 1072
 FT VARSPLIC 1073 1224
 FT CONFLICT 284 284
 FT SEQUENCE 1224 AA; 136186 MW; 572E72129E1401F8 CRC64;
 Query Match 14.8%; Score 1148.5; DB 1; Length 1224;
 Best Local Similarity 28.3%; Pred. No. 8.1e-63;
 Matches 328; Conservative 145; Mismatches 443; Indels 241; Gaps 43;
 2 GRPTRKAPSHSAPLLGLALLRMHQHPRARCPPLCVAGILACGFLLCWGPFSPHQSC 61
 192 GRAAQGSPSHV-----LYKRSTPHAP-----GASEVLVTSRTWELAH----- 230
 62 QALEPOAVSYLSPGAPLK-----GR-----PPSPG-----FQQRQORRAA 99
 231 QPLH-----SSDLRLGLPKQKQHCGRKKYKWPQPKEDLFLPDEYKSLRKHRSLLRSHR 286
 100 GGLHLELVAVGPDVFAQH-QEDTERVLTNLNIGAELELDPISGAQFRVHLNKMVILT 158
 287 NEELNVETLVVDKMMQNHENITVLTILNWSALFKDGTGINIAIVGLILLE 346
 159 EREGAPNITANLSSLLSVCGWSQTINPEDDTPGHADLV--LVITRPDLELPDGNRQVR 216
 347 DRQPGVLVSHADHTLSSFCQWQSLGKMGKQGRTRHDHALLTGLDSCWKNRCPD----TL 402
 217 GVTQLGAGCSPTWSCLTETDPTGLVYTHAIGHSGFLEHGDGAPGSCGSPS-GHVMASD 275
 403 GFAPISGMSKYSRCTINEDTGLAFTIAHESGHNFMIHDG-EGNMCKKSEGNIMSP 461
 276 GAAPRAGLAWSPCSRRLILLSAGRACVWDPPRPQGSAGHPDAPQGLYYSANEQCR 335
 462 LAGNRGVFWSPCSRQYLHKFLSTAQAICLAD--QPKFVKRYKYPEKLPGELYDANTQCK 519

QY 336 VAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLVPLLDGTGCGVKEWKSGBCR 395
 DB 520 WQFGEKAKCMLDFKK-DICKALWCHR---IGRKCTKFWPAABEGTTCGHDMWCRGGQCV 575
 QY 396 SLVELTPTAAVHGRSSWGPSPCSRSSCGGVVTRRQCNPRPAFGGRACVAGDLQAE 455
 DB 576 KYGDEGP-KPTHGHSMDSSWSFPCSTCGGVGSHRSLCTNPKPSHGKFCGSGSTRILKL 634
 QY 456 CNTQACEKTQLEFMSQOCARTDQPLRSSPGGASFYHWAAGAAVPHSQ--GDALCRHMCRAI 513
 DB 635 CMSQKCPRDSVDFRAQAACHNSRRFR-----GRHYKW---KPTQVEDQLCKLYCTAE 586
 QY 514 GESFIMKRGDSFLDTRCMPSGPREDGTLSLCSVSGSCRTFGCDGRMPSQQVWDRQVCGG 573
 DB 687 GDFDFPFLSNKVKDGTFC-----SEDSRNVCDIGICERVGCDNVLSGDAVEDVCGVNG 740
 QY 574 DNSTCSPRKGSFTA-GRAREYVTFVTNLTSTVYIANHRPLFTHLAVR-----IGGR 625
 DB 741 NNSACTIHRGLYTKHHTNQYVHMVTPSGARSIRIYEMNVSTSYISVRNALRYYLNGH 800
 QY 626 YVY--AGKMSISPTTYPSLLEDGRVYRYVALTEDRLPRLEEIRINGPLQEDADIQVYR 683
 DB 801 WTVDMPGRYKFS-GTTF-----DYRRSYNEP-----ENLIATGTNETLIVELLFQ 845
 QY 684 YGEYGNLRRPDITFTYF-----QPKPRQAWVMAVRGCSVSCGAGLRWVNYSCLD 735
 DB 846 G-----RNPGVAWEYSMPRLCTEKQPPAQPSVTNAIVSECSVSCGGQWTVREGCYR 898
 QY 736 QARKELVETVQCGSQPPAWPEACVLEPCPPYVWAGDFGPCASCGGGLRERPVRCVEA 795
 DB 899 DLKFQ-VNMSFCNPKTRPVTLGVPCVSAFPSPWSVGNWSACSRTCGGGAQSRPVQCTRR 957
 QY 796 QSSLKTLTPARCAQAQPAVALETCNPOPCPARWEVSEBSSCTSGAGGAGLENEICV 855
 DB 958 VHYDSEVPASLC-----PQAPS----- 976
 QY 856 PGADGLEAPVTEGPGSVDEKLPAPPCVGMGPGWGHLDATSAKEKAPSPWMSIRTGAQ 915
 DB 977 -----SRQACNSOSCPPAW-----SAG-----PW----- 995
 QY 916 AAHVWTPAAGSCSVSGRGLMELRFLCMD---SALRVVPQBELCGLASKPSRRRVCQAV 972
 DB 996 -----AECSTCGKWRKBAVACKSTNPSARAQLLPDAVCTSEKPR-RMHEACLQ 1045
 QY 973 PC--PARWQYKLA---CSVSCGRGVVRRILYCARAHGEDDGEIILLDTQOGLRPP--E 1025
 DB 1046 RCHKPKKLQWLVSAGSCSVTCERGTQKRFKCAEKVSGKYRE-LASKKCSHLPKPSLE 1104
 QY 1026 PQEACSLBPCPPRWKVMSLQ-----CSASGCLGTARRSVACVQLDQGDV 1073
 DB 1105 LERACAPLPCPRHPPFAAAGFSRGSWFPASWQSQTASCGGVQTRSVQC--LAGRPA-- 1160
 QY 1074 DEACAALVRPEASVPC 1090
 DB 1161 --SGCLLHOKPSASLAC 1175
 RESULT 5
 AT12_HUMAN
 ID AT12_HUMAN STANDARD; PRT; 1593 AA.
 AC P58397;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-12 precursor (SC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
 GN ADAMTS12.
 OS Homo sapiens (Human).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Fetal lung; MEDLINE=21264577; PubMed=11279086;
RX Cal S. Arguella J.M., Fernandez P.L., Lopez-Otin C.;
PA "Identification, characterization, and intracellular processing of
RT ADAM-TS12, a novel human disintegrin with a complex structural
RT organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940(2001).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC expressed in gastric carcinomas and in cancer cells of diverse
CC origin.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase.
CC -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING
CC TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE
CC METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND
CC THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1
CC DOMAINS.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 8 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AJ250725; CAC20419.1; --
CC Genew: HGNC:14605; ADAMTS12.
CC MIM: 606184; --
DR InterPro: IPR001762; disintegrin.
DR InterPro: IPR001818; Pept M10A M12B.
DR InterPro: IPR006025; Pept M Zn BS.
DR InterPro: IPR001590; Peptidase M12B.
DR InterPro: IPR002870; Peptidase_M12B_N.
DR InterPro: IPR000884; TSP1.
DR InterPro: IPR008085; TSP 1.
DR Pfam: PF01562; Pep M12B propep; 1.
DR Pfam: PF01421; Reptolysin; 1.
DR Pfam: PF00090; tsep_1; 6.
DR PRINTS: PR01705; TSP1REPPEAT.
DR SMART: SMC0209; TSP1; 8.
DR PROSITE: PS0215; ADAM MEPRO; 1.
DR PROSITE: PS00546; CYSTEINE SWITCH; FALSE NEG.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE: PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE: PS05090; PLAC; 1.
DR PROSITE: PS00092; TSP1; 6.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 25
FT PROPEP 26 240
FT CHAIN 241 1593
FT DOMAIN 241 464
FT DOMAIN 465 544
FT DOMAIN 542 597
FT DOMAIN 597 700
FT DOMAIN 701 826
FT DOMAIN 823 882
FT DOMAIN 886 942
FT DOMAIN 943 996
FT DOMAIN 996 1315
FT DOMAIN 1312 1365
FT DOMAIN 1367 1421

DOMAIN	1422	1470	TSP TYPE-1 7.
FT DOMAIN	1471	1531	TSP TYPE-1 8.
FT DOMAIN	1534	1574	PLAC.
FT SITE	302	305	POLY-GLU.
FT SITE	208	208	CYSTEINE SWITCH (POTENTIAL).
FT METAL	392	392	ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE	393	393	BY SIMILARITY.
FT METAL	396	396	ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL	402	402	ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD	105	105	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	125	125	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	215	215	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	485	485	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	685	685	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	790	790	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	951	951	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1104	1104	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1275	1275	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1300	1300	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1320	1320	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1371	1371	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1378	1378	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	1503	1503	N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE	1593	1593	AA; 177545 MW; 07P9F48E63D83A3 CRC64;
Query Match	14.5%;	Score 1126;	DB 1; Length 1593;
Best Local Similarity	25.3%;	Pred. No. 2.6e-61;	
Matches	332;	Conservative 156;	Mismatches 501; Indels 322; Gaps 42;
QY	92	RQRRAAGGILHLELLVAVGPDVQAH-QEDTRYVLTNINIGALLRDPGLGAQFRVH	150
DB	235	RSLSRSISIKRWVETLVADTKMIEYHGSNVEYSILTIMMVTGLFHNPSIGNAHIV	294
QY	151	LVMVILTEGAPNITANITSLLSVCGHSQTNPEDDTPGHADLVLYIRFDELPLD	210
DB	295	VVRLILLEEEQGLKIVHAEKTLSSFCWKQKSNPKSDLPVHVDVAVLLTKDI-CAG	353
QY	211	GNR--QVRGVTLGGACSPWCLITDTGDLGVITAEHCHSGLEHGDAGPGSCGFS	268
DB	354	FNRPCTELGLSHLGMQPHRSNINEDSGPLAFTIAHELCHSGFIQHDGKE-NDCEPV	412
QY	269	G---HVMASDGAAPRAGLAWSPCSRRQLLSLSAGRARCVMDPPRPPQGSAGHPDAPG	325
DB	413	GRUPYIMSRQLQYDPTPLTWSKCEEYITRDLRGWGFCLDDIPKKKGLKS--KVIA	469
QY	326	LYYSANEQCRVAFQPKAVACTFAEHLDMQALSCHTDPLDQSSCSRLVPLLDGTECGV	395
DB	470	VIVDHHQCOLQYGPNA---TFCQEVENVQTLWCSV---KGFCSKLDAADGTCQGE	522
QY	386	EKWSKGRCSRLVELTPIAAVHGWSNWSGPRSPCSRSCGGVTVTRRRQNNRPAPFGRA	445
DB	523	KWCMACKCIT-VGKKP-ESIPGGMGRWSPWSHSCRTCGAGVQSAERLCNNPEPKFGKY	580
QY	446	CVGADLQAEKMTQACEKTQLEFMSQQCARIDGQPLRSPPGGSFYHW---GAAVPHSQ	501
DB	581	CTGERKRYLCNVHPCHSEAPTRMQQSEPTVYKN-----ELYHWFPIFNPAPH---	632
QY	502	GDALCRHMCRAIGESFIMKRGDSFLDGTTCMPSPGPREDTLSLCSVSGSRTFCDCGRMDS	561
DB	633	---CELYCRPIDGQFSEKMLDAVIDGTPCEBGG---NSRNVCIINGICKWVCDEYIDS	684
QY	562	QQVMDRCQVCGDNSTCSPRKGSTAGARAREVVTFTVTPNLTSTVYIANHRPFLTHAVR	621
DB	685	NATEDRCGCVCLGDGSSCQTVARMKFQKEGSGVYDIGLIPKGARDIRWMEIEGAGNFLAIR	744
QY	622	IGG--RYVYVAGKMSISPNITPYSLLDGRVEYRVALTEDRLPRELEIRIWGPLEQADIQ	679
DB	745	SEDPKYLNGGFIQWNGNYK--LAGTVFQY-----DRKGDLEKMTGTPNESWVQ	796
QY	680	VYRYGVEYGNLTRPDITFTYFQPKP-----RQAWVAAVR-GPCSVSCAGLRWVNS	732
DB	797	LLFQ-----VTNPGIKYEYTIQKGLDNDVQYFWQYGHWTCSVTCGTGIRRTAH	849

```
QY 733 CLDQARKELVETVQCQSQPPFAPWACVLEPCPPYWAQVDFGFCPSACG-GGLRERFVR 791
D 850 CIKKGR-GWVATFCDEPQNGRQKKEKACPPRWAGEWEACATCGPHGEKRTVL 908
QY 792 CUEAGSLKTLTPPARCRAGAQAFALETCNQ-PCPARWEVSEPSCTAGAGLALE 850
D 909 CIQTWVSDQALPPTDCQ-HLLKPKILL-SCNRDILCPDWTGVNWBECVSCGGVIR 966
QY 851 NCTCPVGD-GLB-----APVTEGFGVDEKLPAPRPPC 882
D 967 SVTCAKNHDEPCDVTKPKNSRALCGLQCCPSRRVLKPNKGTISNGKNPPLK-PVPPT 1025
QY 883 VG---MSCPPGHLDATSAGEKAPSPWGSIRTAQAHAHWTTPAA-----GS 926
D 1026 SPRMLTTTGPESMSTSPAISSPSTASKEDLGKQWQDSSTQBELSRVLIISGS 1085
QY 927 CS-----VSCQGL----- 935
D 1086 TSQILTQSLSIQPSEENVSSDTGPTSEGLVATTSGSLSSRRNPITWPTVPTNT 1145
QY 936 -----MELRFLCMD-----SALRVP----- 950
D 1146 LTKGPEMEIHSGSEBEREQPKDESNFVWTIRVPCNDAPVSTEMPLAPLTPDLNR 1205
QY 951 -----VQELC-----GLAKPGSRREVQAVPCPA-----R 977
D 1206 ESWWPPFSTVMEGLLFSQRTTSETGPRVEGMVTEPANTLLPLGDHQPEPSGKTNR 1265
QY 978 WQYKL-----AACS 986
D 1266 NHLKLPNNNQTKSSEPVLTFEEDATSLITBGLFLNASNYKLTNGHGSAAHIVGNWSSCS 1325
QY 987 VSCGRGVRRILYCARAHGSDDBEILLDTCCQGLPRPEQACSLBPCPPRWKMSLGP 1046
D 1326 TTCCGLGAYWKBCTTQ-----MDSDCAIQRPDPAPKRLRCP-AGWKVGNWSK 1374
QY 1047 CSASCGLGATARRSVACVQLDQGVDEDEACALVPEASVPCLIADCTVYRHHVGTWE 1106
D 1375 CSRNCGGFKIREIQCVDSRDRNLFHFCQFLAGIPPLISMSNPEPCE-AWQVWPMSQ 1433
QY 1107 CSVSCGDGQRRRDTCLGPOQAQVPAFDFQHLPKPVTVRC-----WA 1150
D 1434 CSRSCGGVQVERGVC-----PGLCDWTXRTSTWMSCNHEHLCCHNA 1475

RESULT 6
AT18 HUMAN
ID AT18 HUMAN STANDARD; PRT; 1081 AA.
AC Q8TE60;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-18 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 18) (ADAM-TS 18) (ADAM-TS18).
GN ADAMTS18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney
CC and in adult brain, prostate, submaxillary gland, and endothelium.
```

```
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ311903; CAC83612.1; .
CC Genew; HGNC:17110; ADAMTS18.
CC MIM; 607512; .
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR000885; TSP1.
CC Pfam; PF01562; Pep M12B_Prosep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; TSP1; 4.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS02115; ADAM_MEPRO; 1.
CC PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS00142; ZINC PROTEASE; FALSE NEG.
CC PROSITE; PS00142; ZINC PROTEASE; FALSE NEG.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix.
CC SIGNAL; 1 47 POTENTIAL.
CC PROPEP 48 284 BY SIMILARITY.
CC CHAIN 285 1081 ADAMTS-18.
CC DOMAIN 285 497 METALLOPROTEASE.
CC DOMAIN 498 577 DISINTEGRIN-LIKE.
CC DOMAIN 588 643 TSP TYPE-1 1.
CC DOMAIN 644 749 CYS-RICH.
CC DOMAIN 750 876 SPACER.
CC DOMAIN 877 931 TSP TYPE-1 2.
CC DOMAIN 933 991 TSP TYPE-1 3.
CC DOMAIN 993 1048 TSP TYPE-1 4.
CC SITE 254 254 CYSTEINE SWITCH (POTENTIAL).
CC METAL 435 435 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT_SITE 436 436 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 439 439 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 445 445 ZINC (CATALYTIC) (BY SIMILARITY).
CC CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 313 313 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 744 744 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 837 837 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 908 908 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 1081 AA; 119655 MW; 0438BF645676461E CRC64;

Query Match 13.4%; Score 1040.5; DB 1; Length 1081;
Best Local Similarity 28.6%; Pred. No. 2.9e-56;
Matches 275; Conservative 119; Mismatches 355; Indels 213; Gaps 32;

QY 94 QRRAAGGI---LHLELLVAVGDDVFOAH-QEDTERVYLTWLNIGAEILLRDSLGAQFRV 149
D 281 RFRSAGKSQGLNVETVAVDKKWKVKGKGNVTYLTVMKVSQ-LFKDGTIGSDINV 339
QY 150 HLKVMVLTPEGAPNITANLTSLLSVCGWSQINPEDDTPDGHADLVLYITRFDL-EL 208
D 340 VVVSLLLEQEPGLLINHADOSLNSFCQMSALIGNKGRHDA---ILLTGFDICSW 396
```


QY	209	PDGNRQVRGVTQLGGACSPWWSCLITDFTDLGVTTIAIHGSHFGLHDGAPGSCCGPS	268
DB	397	KNEPCDTLGFAPTSNGMSCKYRSCTINEDTGLGUAFTIAHESGNFGMIHDG-EGNPCRKA	455
QY	269	-GHVMSADGAAPRAGLAWSPCSRRQLLSLLSAGRCVMDPPRPQPSGAGHPDPAQPLGY	327
DB	456	EGNINSEFTLGNNGVFSWSSCSQYLKKFLLSTFQAGCLVDEPK-QAGQYKY-PDKLPGQI	513
QY	328	YSANEQCRVAFGPKAVACT--FAREHLDMQALSCHTDPLDQSSCRLLVPLLDGTECGV	385
DB	514	YDADTCKMGFGAKAKLCSLGFVK--DICKSLWCHR--VGRHCETKFMPAAEGTVCGL	567
QY	386	EKWCSGKRCRSLVELPTPIAAVHGHWGWSGPRSPCSRSCGGVYTRRQCNPPAPGGRRA	445
DB	568	SMNCRGQCQVKFGELGP-REIHQWSAWSKWSCESTRCTCGGKVFQERHCNNPKPYGGIF	626
QY	446	CVGADLQAEKNCNTQACEKTOLEFMSQOCARTDGOPLRSFPGSASFYHWGAAPVHSGD DAL	505
DB	627	CPGSSRIYQCNINPCNENSLDFRAQCAEYNSKPFK-----GWFWQWKYTKVEEDR-	680
QY	506	CRMCMRAIGESFIMKRGDSFLDGTTCRMPSPGREDGTLSLCSVSGSCRTFGCDGRMDSQQVM	565
DB	681	CKLYCKRAENFEFFAMSGKVKDGTTPCSPN--KND-----VCIDGVCELVGDHGLGSKAVS	734
QY	566	DRQVCGGDNSTCSPRKGSF-TGABREYVTFLLVTNLTSVYIAHRPLPFTHLAVR-IG	623
DB	735	DACGVCKGDNSTCKFYKGLYNHOKANEYFPVVLIPAGARSIEIQELQVSSSYLAVRSLS	794
QY	624	GRYVWAKMISIPNTTYPSSLLEDGRVRYVALTEDRLPRLEETIRMGFLOEDADIOVYRR	683
DB	795	QKYLTLGGWSDWPGEPF--FAGTTPEYQKSFN-----RPERLYAPQPTNLTLVFELMQ	847
QY	684	YGBEYGNLTPRDIPTFTYFQPK-----PQAWVWAAVRGFCVSVCAGLRWVNYSC L	734
DB	848	-GK-----NPGIAWKYALPKVWNGTPTPKRPAYTWSIVQSECVSCGGYINVKALCL	900
QY	735	DQARKELVETVQCGSQOPPAWPEACVLEPCPPYVAVGDFGFCPSACSCGGGLRERPRVCVE	794
DB	901	RDQNTQ-VNSSFCSAKTKPVTPEKICNFAFCSPAYWMPGEWSTCSKSCAGGQQRKIQCVQ	959
QY	795	AQGSLLKTLPPARCEAGAQPAVALETNCNPQPCPARVEVSEPSSTCSAGGAGLALENETC	854
DB	960	-----KKPFQKEAVLHSLCP-----VSTPTQV-----	982
QY	855	VPGADGLEAPVTEPGSGVDEKLPAPFCVGMGSCPPGPHGLDATSAGEKAPSPWGSINTGA	914
DB	983	-----QACNSHACPPQW-----SLG-----PW-----	999
QY	915	QAAHVWTPAAGSCSVSCGRGLMELRFLCMOSALRPVVOELGCLASKPGSRREVQCAVPC	974
DB	1000	-----SQCSKTCGRGVRKRELLCKGSA-----	1021
QY	975	PARWOYKLAACSVSCGRVVRRIIYCARAHGEDDGEIILLDTQCGGIPREPQEQACSLP	1034
DB	1022	-----AETLPESQCTSLRPELOEGCVLGR	1046
QY	1035	CP 1036	
DB	1047	CP 1048	

RESULT 7

AT10 HUMAN

AT10 HUMAN

STANDARD;

PRT; 1077 AA.

AC Q9H324

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-10 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).

GN ADAMTS10.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

QY	13.2%; Score 1026; DB 1; Length 1077;	975	PARVOYKLAACSVSCGRVVRILYCARAHGEDDEIILDTQCGLPRPPEQACSLP	1034
Db	Best Local Similarity 26.0%; Pred. No. 2.3e-55;	966	PATMR-----CNLRR	975
	Matches 305; Conservative 126; Mismatches 425; Indels 316; Gaps 41;			
QY	2 GRPTKAPSHAPLGLALLRHQHRPARCPPLCVAGILACGLGCGWSPFQSC- 60	1035	CP-RWKWSLGPSCSACGLTARRSVACVQLDQGVDEVDAAALVRDEASVPLIA	1093
Db	137 GGPGRSRSPESSGP--HVVYKRSLSRHP-----HLDACG 169	976	CPARWVAGECSACQGVQQRVCRTS-HTGQ--ASHECTEALRPP----- 1022	
QY	61 -----LQALEPQAVSSYLSPGAPLKRPP--SPGFQRCQRORRAAGILHL 105	1094	DCYRWHVGTWMECSVSC-----GDGIQRRRD 1120	
Db	170 VRDEPKWGRPWMLRTKP-----PPAPLNETERGQGLKRSVSRER-----YV 215	1023	-----TTOCEAKCDSFTPGDGPFECKD 1045	
QY	106 ELLVAVGPVFOAH-QEDTERYVLTNINCAELRLDPSLAQAFVHLVGMVILTEPEGAP 164			
Db	216 ETLVADKMWAYHGRDVEQYVLAIMNIVAKLFQDSSLGSTVNILVTRLLILLTDQPL 275			
QY	165 NITANLTSSLVCGWSQTI-----NPEDDTPGHADLVLYITRFDL-----ELPDGMR 213			
Db	276 EITHAGKSLDSFCKWQKSVNHSGHGNAIPENGVAHDTAVLITRDIYCIYKKNPCG- 333			
QY	214 QVRGVTQLGACSPWSCILTEGTDLGVITIAHIGHSFLEHGDGAPGSCGSGH--- 270			
Db	334 -TLGLAPVGMGERERSCSVNEDIGLPQAFITIAHIGHTFGMNDGV-GNSCGARGQDPA 391			
QY	271 -VNASDGAAPRAGLWSPCSRRLSLLSAGARCVMDPPRPPQSAGHPDPAQGLYYS 329			
Db	392 KUMAAHITWKTNPFWSSCNRDYITFLDGLGLCLNNRP---PRQFVYPTVAPQAYD 448			
QY	330 ANEQCRVAFGRKAVACTFAREHLDMCOALSCHTDPLDQSS--CSRLVPLLDGTEC---GV 385			
Db	449 ADEQCRFQHGVRKQCKYG---EVSSELWC---LSKSNRCITNSIPAAGTLCQHTI 500			
QY	386 EK-WCKGKCRSLVELTPIAA-----VHGRWSMGPRSPCRSCGGVYVTRRQNNRPPA 440			
Db	501 DKXWCYKRVK-----VPFGRPEGVDGAGPWTGDCSRTCCGGVSSSRHCDSPRPT 554			
QY	441 FGRACVGLADLAQEMCMTQACEKTOLEFMSQOQARTDQPLRSPGASFYHWAANVPHS 500			
Db	555 IGKCYCLGERRHRSNCTDDCPGQSDRFVQCSFDSIPPR-----GKFKYK--KTYRG 607			
QY	501 QDCLCHMCRAIGESFIMKRGDSFLDGTGCMPSGPREDTGLSLCVSGSCTRFTGCDGRMD 560			
Db	608 GGVKACSLTSLAGFNFYTERAAAVDGTGCRPD-----TVDLCVSGECKVQCDRLVG 661			
QY	561 SOQVMDRCVCGDNDSTCSRKGSFT-AGRAREYVTELTTPNLTSLVYIANHRLPFLTHA 619			
Db	662 SLDREDKCRVCGDGSACETIEGVFSPASPGAGYEDVWIPKGSVHIFIQDLNLSLSLA 721			
QY	620 VIRGRYVAGKMSISPNITYPSLLEDGRVVEYVALTEDRLPRLEIRIMWPLEQDADIQ 679			
Db	722 LKGDQESLILLEG- --PGTPQHRLPLAGATTFLRQGPQDQVQSLEAL--GPNASLIVM 775			
QY	680 VYRYGEEYGNLTPDITFTYFOP-----KPRQWVAWVRGPGSVSCAGLRWVWVYSL 734			
Db	776 VLAR-----TLPALRYAFNAPIARDSLPFYSWHYAPWT-KCSAQAGGSQVQAVECR 827			
QY	735 DOARKELVETVOGSOQOPPAWPEACVLEPCPPYAVGDFGPGCSACGGGLRERPVRCVE 794			
Db	828 NQLDSAVAPHYCSAHSKLEPRQACNTEPCPPDWVGNWLSRSCDAGVRSRVVC- 885			
QY	795 AQSGLLKTLPAPRCRAGAPAVALETCNPPCFARWEVSEPPSCTSAGGAGLALNETC 854			
Db	886 -----QRRVSAEKEALDDSDAC-QQRP----- 907			
QY	855 VPGADGLEAPVTEGPGSVDEKLPAPEPCVGMSCPPGWHGLDATSAGEKAPFWGSIRTGA 914			
Db	908 -----PVLEACHGPTCTCPPEWAALD----- 926			
QY	915 QAAHVWTPAAGSCVSCGRGLMELRFLCMDSALRPVQVEELCGLASKPGSRREVCOAVPC 974			
Db	927 -----WS-----ECTPSCGGLRRVWLCKADHATLPPAHCSPAAPK----- 965			

RESULT 8

AT17 HUMAN	AT17 HUMAN	STANDARD;	PRT; 1095 AA.
AC	Q8TE56;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	15-MAR-2004 (Rel. 43, Last annotation update)		
DE	ADAMTS-17 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 17) (ADAM-TS 17) (ADAM-TS17).		
GN	ADAMTS17.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21856482; PubMed=11867212;		
RA	Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,		
RA	Lopez-Otin C.;		
RT	"Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";		
RL	Gene 283:49-62(2002).		
CC	!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).		
CC	!- TISSUE SPECIFICITY: Expressed in fetal lung, in adult brain, prostate, and liver.		
CC	!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).		
CC	!- SIMILARITY: Belongs to peptidase family M12B.		
CC	!- SIMILARITY: Contains 1 disintegrin-like domain.		
CC	!- SIMILARITY: Contains 1 TSP domain.		
CC	!- SIMILARITY: Contains 5 TSP type-1 domains.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).		
CC	EMBL; AJ315735; CAC86016.1; -		
DR	MEROPS; M12.027; -		
DR	Genew; HGNC:17109; ADAMTS17.		
DR	MIM; 607511; -		
DR	InterPro; IPR001762; Disintegrin.		
DR	InterPro; IPR001818; Pept M10A_M12B.		
DR	InterPro; IPR006025; Pept M_Zn_BS.		
DR	InterPro; IPR001590; Peptidase M12B.		
DR	InterPro; IPR002870; Peptidase_M12B_N.		
DR	InterPro; IPR000884; TSP1.		
DR	InterPro; IPR008085; TSP_1.		
DR	Pfam; PF01562; Pep M12B_Propesp; 1.		
DR	Pfam; PF01421; Reprolysin; 1.		
DR	Pfam; PF00090; tsp_1; 5.		
DR	PRINTS; PR01705; TSPIREPEAT.		
DR	SMART; SM00209; TSP1; 5.		
DR	PROSITE; PS00215; ADAM_MEROPS; 1.		
DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.		

DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0900; PLAC; 1.
DR PROSITE; PS50092; TSP1; 5.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 27 POTENTIAL.
FT PROPEP 28 223 BY SIMILARITY.
FT CHAIN 224 1095 ADAMTS-17.
FT DOMAIN 224 452 METALLOPROTEASE.
FT DOMAIN 453 542 DISINTEGRIN-LIKE.
FT DOMAIN 543 598 TSP TYPE-1 1.
FT DOMAIN 599 701 CYS-RICH.
FT DOMAIN 702 779 SPACER.
FT DOMAIN 800 860 TSP TYPE-1 2.
FT DOMAIN 861 922 TSP TYPE-1 3.
FT DOMAIN 925 968 TSP TYPE-1 4.
FT DOMAIN 972 1029 TSP TYPE-1 5.
FT DOMAIN 1045 1084 PLAC.
FT DOMAIN 60 120 AEG-RICH.
FT SITE 201 201 CYSTEINE SWITCH (POTENTIAL).
FT METAL 389 389 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 390 390 BY SIMILARITY.
FT METAL 393 393 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 399 399 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 167 167 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 483 483 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 785 785 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 839 839 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;
Query Match 13.0%; Score 1005; DB 1; Length 1095;
Best Local Similarity 27.5%; Pred. No. 4 5e-54;
Matches 318; Conservative 114; Mismatches 409; Indels 316; Gaps 45;
QY 4 PTPKKA-PESHAPLLGLAL-----LRMHQR-----HPARPPLC--- 37
DB 69 PAAPRAPGERALLHLPAFRDLYQLRRDLFLSRGFEVEBAGARRRGPAELCFYS 128
QY 38 -----VAGILACGFLLCGWPSPHFQSCLOALEPQAVSVLSGAPLKR----- 82
DB 129 GRVLGHPGSLVSLSACGAGGLVGLQLGQ-----EQVLIQPLNNSGGPFGSREHLIRR 182
QY 83 ----PPSPGFQRQ-----RORRAAGGLH---LELLVAVGP 113
DB 183 KSLTSPSAEAQRPEQLCKVLTEKKTWGRSPDRWRERNRNLRTSEHTVELLVADA 242
QY 114 DYFQAH-QEDTERYVLTNLNIGALLRDPDLGQAFRVHLVYKMWILTEPEGAPNITANLTS 172
DB 243 DMVQYHGABAAQRFILTVNMVYMFQHSGLGKINQVTKLVLLRQPAKLSIGHGER 302
QY 173 SLLSVCGW-----SQTNIPEDTDPGHADLVLYTRFDEL-PGNNRQVRGVTQ 220
DB 303 SLESFCHWQNEYEGGARYLGNQVPGGKDDPLVDAAVFVTRTDFCVHKDEPCTVTGIAY 362
QY 221 LGGACSPHSCLTEDTGDPLGVTHAIEIGHSGLEHDGAPGSCGSGHVMASDGAAPR 280
DB 363 LGGVCVSAKRCVLAEDNGLNLAFTHAELGHNLMNHDD-DHSSCAGSRHIMSGEWKVR 421
QY 281 --AGLAWSPCSRQLLSLSAGRARCVW-DPPRPQPGSAGHP---PDAQPLGLYISANEQC 334
DB 422 NPDSLWSWSSCRDLDLENFLSKVSTCLLVTDPSQ-----HTVELPHKLPGMHVSANEQC 476
QY 335 RVAFGPKAVACTFAR--EHLDMQALSCHTDPLDQSSCSLLVPLLDGTEGCVKWCCKG 392
DB 477 QILFGMNA---TFCRNNEHL-MCAGLMCLVE--GDTCKTKLDPLDGTGCGADKWCRCAG 530
QY 393 RCRLSLVELTPIAA-VHGRWSWSPRSPCSRSCGGGVVTRRQCNPRPAFGRCACVADL 451

DB 531 EC---VSKTPIPEHVDGWSWPGAWSMSCRTCTGTGARFRQRKCDNPPPGGTHCPGASV 587
QY 452 QAEMCNTQACEKTLQLEFMSQQCARTDGGQLRSPGGASFYHGAAYVPHSGDALCRHMC 511
DB 588 EHAVCENLPCPKGLPSFDQCCQAH-----RLSPKKK-----GLLTAVVVDDPECELYCS 638
QY 512 AIGSFIMKRGDSFLDTRCMPSPGREDGTLCLVSGSCTRTFCGDRMDSQQWDRQVC 571
DB 639 PLGKESPLLAVDRVLDTGTPC---GPYE---TDLGVHGKQKICGCDGLIGSAAKEDRCGVC 692
QY 572 GGDNSTCSPRKGSFTAGRAREYVTLTVPNLTSVYIANRPLFTHLAVRIGRYVAVGK 631
DB 693 SGDGTCHLVKRGDFSHAR-----GTLAKDSGK 719
QY 632 MSISPNNTYPSLLEDRV---EYRVALTEDRLPR---LEEIRINGLQBDADIVYVRY 684
DB 720 CSIN-----SDWKTELPGEFQIAGTVYVRGLWEKISAKGPTKLPPLHMLVLLFH 770
QY 685 GEEYGNLTPDPITFTYFQP-----KPRQA---VWMAVRGPCVSCGAGLRWV 729
DB 771 DDYVG-----IHVEYTVVNRTAENQSEPEKPDQSLFIWTHSGWEG-CSVQCGGERRT 823
QY 730 NYSCLDQARK--ELVETVQCQSQQPPAPWPEACVLEPCPPYVAVGDFGPGSASCGGLRE 787
DB 824 IVSCTRIYVNTKTLVNDSDCPQASRPQVRRCNLHPCQSRWVAGPWPSPCSATCEKGFQ 883
QY 788 RPVRGV-BAQGSLLKTLTPARCAGAAQPAVALETCPNCPARWEVSEPSSTCSAGGAG 846
DB 884 REVTCVYQLQNGTHVATRYLYC--PGRPFA-AVQSCGQDCLSIWEASE----- 929
QY 847 LALENETCPVAGDGLAEPVTEGPGSVDEKLPAPEPCVGMSCPPGWHGLDATSAGEKAPSP 906
DB 930 ----- 929
QY 907 WGSIRTGAAAHVWTPAAGSCSVSCGRGLMELRFLWDSALRVPVQBELGLASKPGSRR 966
DB 930 -----WS-----QCSASCGKGVKRTVACTNS-----QCKDASTRPA-E 964
QY 967 EVCQAVPCPARQY-KLAACSVSCGEGVVRILYCAR---AHGEDDGEILLDTCCQGL 1021
DB 965 EACEDYSGCYEYKTDGWSCTCSSTCKGLOSRRVQCMKVTGRHG-----SECPAL 1014
QY 1022 PRPEQEAACSLPFCPPR 1038
DB 1015 SKPAPYQCYQEVNDR 1031
RESULT 9
ID AT52_MOUSE STANDARD; PRT; 1213 AA.
AC Q8C9W3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2) (Procollagen I/II
DE amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC
DE I-NP) (Procollagen N-endopeptidase) (pNPI).
GN ADAMTS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bratt D., Bruscia V., Chothia C., Corbani L.E., Cousins S.,

RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinchik S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konegaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteau G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sadelin A., Schneider C., Sample C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kogawa I.,
RA Miyazaki A., Sakai K., Sakaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-FVB/N; TISSUE-Breast tumor;
RX MEDLINE=22388557; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kretzmar M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gough J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
CC to fibril assembly. Does not act on type III collagen. May also
CC play a role in development that is independent of its role in
CC collagen biosynthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC at Ala1(I) at Pro1-Gln and of alpha1(I) and alpha2(I) chains
CC at Ala1-Gln.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to
CC collagen type XIV (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL; AK040370; BAC30572.1; -;
CC DR EMBL; BC048456; ABA46456.1; -;
CC DR MGI; MGI:1347356; Adamts2.
CC DR InterPro; IPR001762; Disintegrin.
CC DR InterPro; IPR006025; Pept M.Zn BS.
CC DR InterPro; IPR001590; Peptidase M12B.
CC DR InterPro; IPR002870; Peptidase_M12B_N.
CC DR InterPro; IPR000884; TSP1.
CC DR Pfam; PF01562; PEP_M12B_propep; 1.
CC DR Pfam; PF01421; Reprolysin; 1.
CC DR Pfam; PF00090; TSP_1; 4.
CC DR SMART; SM00209; TSP1; 4.
CC DR PROSITE; PS0215; ADAM_MBP; 1.
CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
CC DR PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
CC DR PROSITE; PS0900; PLAC; 1.
CC DR PROSITE; PS0092; TSP1; 4.
CC DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
CC KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Collagen degradation; Extracellular matrix.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 260 BY SIMILARITY.
FT CHAIN 29 1213 ADAMTS-2.
FT DOMAIN 261 471 METALLOPROTEASE.
FT DOMAIN 480 560 DISINTEGRIN-LIKE.
FT DOMAIN 561 617 TSP TYPE-1 1.
FT DOMAIN 618 722 CYS-RICH.
FT DOMAIN 723 851 SPACER.
FT DOMAIN 855 913 TSP TYPE-1 2.
FT DOMAIN 915 975 TSP TYPE-1 3.
FT DOMAIN 976 1030 TSP TYPE-1 4.
FT DOMAIN 1060 1098 PLAC.
FT SITE 692 694 CELL ATTACHMENT SITE (POTENTIAL).
FT METAL 409 409 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 410 410 BY SIMILARITY.
FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 419 419 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 252 252 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 949 949 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 950 950 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 994 994 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1032 1032 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1099 1099 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1147 1147 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1152 1152 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 1213 AA; 135298 MW; B27431E00443EDB5 CRC64;
Query Match 12.6%; Score 978; DB 1; Length 1213;
Best Local Similarity 27.7%; Pred. No. 2.3e-52;
Matches 301; Conservative 133; Mismatches 383; Indels 268; Gaps 47;
QY 24 MHQRHPRAPCPICVAGILACGLFLLGCGWSPHFOOSCLQALEPQAVSSVLSG----- 76
Db 210 VYRRPPTPKPPVPS-----EPQALDTGVSGQNLDSLRS 242
QY 77 --APLKRRPPSPGQROQRQRRAGGILHLELLVAVGPDYQAH-QEDTERVLTNLNI 133
Db 243 ALGVLEERINS-----SRRRVRHATDDDDYNTFVLLGVDDSVVQVFGKHQVLYLLTMNI 298
QY 134 GAELLRDPSPSLGAOFRRVHLVWVILTEPEGADNI-TANLTSSLLSVCGWSSQ-TINPEDD 191
Db 299 VNEIYHDES LGAHINNVVLRILLSHAKSMELIIGNFSQSLNVCRWAYLQCKPTDHD 358
QY 192 PGHADLVLYITREDFLEPDGNRQVRGVTQLGGACSPWTSCLITEFDPLGVITAEIGH 251
Db 359 EYH-DHAIFLTRQDF-----GPSGMQGVAPVTGMCHPVRSCTLNHEDGFSSAFVWAHETGH 413
QY 252 SFGLGHDGAPSGCGPS---GHVMSADGAPRAGLAWFCSRRLQLLSAGARCVNDP 308

Db 414 VLGMEDHG-QGNRCGDEVRLGSIAPLVQAFPHFRFWSRCSQQELSRYLHS-----YDC 466
Qy 309 PRPGSAGHPDQA-PGLYYSANEQCRVAFGKAVACTFAREHLDMCQALSCHTDLDQ 367
Db 467 LRDDFFAHDWAPALPQLPGLHYSMNEQCRDFGLGMMCTAFRT-FDPCCKQLWC-SHPDNP 524
Qy 368 SSCSLLVPLLDGTEGVEKWKSGKGRSLVELTP-TAAVHGWSNGRSPCSRSGGG 426
Db 525 YFCKTKGPPLDGTCAPGKCHFGHC---IWTPDILKRDGNWGAWTFPGSCSRCTGTG 581
Qy 427 VYTRERQCNPRPAFGGRACVADLOAEMCNTOQACEKTQLEFMSQQCAQTIDGQLRSPG 486
Db 582 VKFRTRQCNPNPANGRCISGLAYDQPCNSLADFREQCQOWD---LYPEHG 638
Qy 487 GASFYHWGAAPHSGDA--LCHMC--RAIGSFIMKRGDSFLDGTCTCMPSGPREDTGL 542
Db 639 DVQ-HHW---LPHEHRDAXERCHLYCSKETGEVVSMMKR--WVHDGTRC-----SYKDAF 687
Qy 543 SLVSGSCRTFGCDGRMSQVWDRCOVCGDNSTCSPKGSFT-AGRARVYVFLTVP 601
Db 688 SLVCRGDCRVGCDGVIGSRKQEDKCGVCGDNTHCKVWGTFTSRPKQDYIMKFEIPA 747
Qy 602 NLTSVVIANHRPLFTHLAVR--IGRRVVAGKMSISNTTPTPSLLEDG-RVEYRVALTED 658
Db 748 GARHLDIQEADTTSHLSVKNLETGKFLINEENHLDNSR--SFIAMGVEYR---NED 802
Qy 659 LRPLREIRIWPLOEDADIQVRRYGEEXGNLTPDITTFYF-----QP 703
Db 803 ERETLOTI---GPHGTITVLIPE-GD-----TRISLYKYMIEDSLNVDNNVLEDD 853
Qy 704 XPRQAVV---WAAVRGPCSVSCGAGLRWVNSCLDQARKELVETVQCGSQPPAMPBAC 760
Db 854 AVRHEWALKWS---PCSKPCGGSGQFTKYGCRRRLDSKMVHRAFCALAKPKAIRAC 909
Qy 761 VLEPC-PPYVAVGDFGPCSACG--GRLRERVRVE--AGSLLKTLPLPARCBAGAQPAV 817
Db 910 NFQECSPQVWVTGEWBEPTCQSGTQGVRSVCIQPLHNNTTSVHTKGNDRHPE--- 966
Qy 818 ALETQNPQPCPARWEVSEPSSTCSAGGAGLALENCTCPGADGLEAPVTEGPGSVDEKLP 877
Db 967 SRACNRELCPGRW----- 980
Qy 878 APEPCVGMSCPPGWHLDATSAKEKAPSPWGSIRTAQAHHVWTPAAGSCSVSCGRGLME 937
Db 981 -----RAGS-----WS-----QCSVTCNGTQE 998
Qy 938 LRFLCWDALRVPVQBELGLASVGRREVCQAVPCPARWQYKLAACSVSCGRGVVRI 997
Db 999 RPYLCRTADDNFGVCRE-----ERPETAR-ICRLAPCP----- 1030
Qy 998 LYCARAHGEDDGBEILLDTQCGHLPREPQACSLPCPPPRKVMKSLGPCSASCGLGTAR 1057
Db 1031 -----RNGSDPSKSYV---VQWLSRPDPDS-----PIQTSSKQCCQGD----- 1067
Qy 1058 RSVAC 1062
Db 1068 KSMFC 1072

RESULT 10
AT14 HUMAN
ID AT14 HUMAN STANDARD; PRT; 1223 AA.
AC Q8WXS8; Q8TE55; Q8TEY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-14 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 14) (ADAM-TS 14) (ADAM-TS14).
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=21638661; PubMed=11779638;
RA Bolz H., Ramirez A., von Brederlow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
RT metalloproteinase family";
RL Biochim. Biophys. Acta 1522:221-225(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Fetal lung;
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains";
RL Gene 283:49-62(2002).
RN [3]
RP SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND
RP ALTERNATIVE PROMOTER USAGE.
RX MEDLINE=21839041; PubMed=11741898;
RA Collige A., Vandenbergh I., Thiry M., Lambert C.A., Van Beeumen J.,
RA Li S.-W., Prockop D.J., Lapiere C.M., Nussgens B.V.;
RT "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying
RT high homology with ADAMTS-2 and ADAMTS-3";
RL J. Biol. Chem. 277:5756-5766(2002).
CC -!- FUNCTION: Has a aminoprocollagen type I activity processing
CC activity in the absence of ADAMTS2. Seems to be synthesized as a
CC latent enzyme that requires activation to display aminoprocollagen
CC peptidase activity.
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative promoter;
CC Comment=2 isoforms, A (shown here) and B, are produced by use of
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q8WXS8-1; Sequence=Displayed;
CC Name=B;
CC IsoId=Q8WXS8-2; Sequence=VSP_006958;
CC Name=C;
CC IsoId=Q8WXS8-3; Sequence=VSP_006958, VSP_005501;
CC Note=Produced by alternative splicing of isoform B;
CC Name=D;
CC IsoId=Q8WXS8-4; Sequence=VSP_005501;
CC Note=Produced by alternative_splicing of isoform A;
CC -!- TISSUE SPECIFICITY: Expressed in retina and at low levels in
CC brain, lung and placenta. High expression in fetal tissues.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AF358666; AAL40229.1; -
DR ENBL; AJ345098; CAC87943.1; -
DR ENBL; AF366351; AAL79814.1; -
DR MEROPS; M12.024; -
DR Genew; HGNC:14899; ADAMTS14.
DR MIM; 607506; -

```
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01562; Pept_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS02015; ADAM_MPRO; 1.
DR PROSITE; PS00346; CYSTEINE_SWITCH; FALSE NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
DR PROSITE; PS03000; PLAC; 1.
DR PROSITE; PS00092; TSP1; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE NEG.
DR HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Collagen degradation; Repeat; Extracellular matrix;
KW Alternative splicing; Alternative promoter usage.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 252 BY SIMILARITY.
FT CHAIN 253 1223 ADAMTS-14.
FT DOMAIN 253 460 METALLOPROTEASE.
FT DOMAIN 461 551 DISINTEGRIN-LIKE.
FT DOMAIN 552 607 TSP TYPE-1 1.
FT DOMAIN 608 729 CYS-RICH.
FT DOMAIN 730 846 SPACER.
FT DOMAIN 847 907 TSP TYPE-1 2.
FT DOMAIN 908 967 TSP TYPE-1 3.
FT DOMAIN 968 1022 TSP TYPE-1 4.
FT DOMAIN 1059 1097 PLAC.
FT DOMAIN 875 878 POLY-ARG.
FT DOMAIN 1100 1221 PRO-RICH.
FT ACT_SITE 399 399 BY SIMILARITY.
FT METAL 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 941 941 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 68 Missing (in isoform B and isoform C).
FT VARSPLIC 368 368 /FTid=VSP_006958.
FT VARSPLIC 368 368 G -> GMGQ (in isoform C and isoform D).
FT CONFLICT 868 868 /FTid=VSP_005501.
FT CONFLICT 884 884 Q -> R (IN REF. 2).
FT CONFLICT 901 901 Q -> H (IN REF. 2).
FT CONFLICT 923 923 C -> S (IN REF. 2).
FT CONFLICT 1024 1024 C -> Y (IN REF. 2).
FT CONFLICT 1024 1024 N -> S (IN REF. 2).
FT SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F170F CRC64;
Query Match 12.38; Score 956.5; DB 1; Length 1223;
Best Local Similarity 26.88; Pred. No. 4.7e-51;
Matches 310; Conservative 142; Mismatches 412; Indels 291; Gaps 51;
QY 93 QRRRAAGGILHLELLVAVGPDVFQAH-QEDTERVLTNLNIGALLRDPISLGAQFRVHL 151
Db 249 RKRHAKPGSYSLVLLVDDSVRFHGHKEHVQVYLTLMNIVDEIYHDESLSGVHINIAL 308
QY 152 VKRVILTEPEGANI-TANITSSLLSVCQWSQTINPEDTDFGHADLVLYITRFDLPLPD 210
Db 309 VRLIMVGYRQSLIERGNPSRSLQCVRWASHQQRQDPSSHAENHHDVFLTRQDFG-PS 367
QY 211 GNRQVRGVTOLGACSTWNSCLITDGFGLGVTFIAHIGHSFGLHGDGAPSGGCPSS--- 268
Db 368 GYAPVTGM-----CHPLRSCALNHEDGFSSAFIAHETGHVGLNMEHDG-QQNGCADETS 420
QY 269 -GHVWASDGAAPRAGLAWSPCSRQLLSLSAGRACV----WDPERPQPSAGHPDDAQ 323
```

```
Db 421 LGSVVAAPLQAAAFHFRHWSRCSKLELSRYLPS--YDCLLDPPDFPAMFQ-----PPEL- 471
QY 324 PGLYTSANEOCRVAFQPKAVACTFARHBLDMCOALSCHTDPLDQSSCSRLLVPLLDGTTC 383
Db 472 PGINTSMDEQCRDFDGSYGQTC-LAFRTFFPKQLWC-SHPDNFYFCKTKKGPPLDGTTC 529
QY 384 GVEKWCCKGRCSLVELTP--IAAVHGRWSSWGRSPSCSRSCGGGVVTRRRRNNPRPAP 441
Db 530 APGKWCFKGHC---IWKSPQTYGQDGGWSWTWFGSCSRSCGGGVRSRSCNNPSPAY 586
QY 442 GGRACVADLQAEKMCNTQACEKTLQEFMSQOCARTDQQLRASSPGAGFY-HWGAA---V 497
Db 587 GGRPCLGPMFEYQVNCSEECPTGYEDFRAQCAKEN-----SYVYHQAAGKHSV 635
QY 498 PHSQGD--ALCRHMCRAIGESFIMKRGDSFLDGRCPMPGPRPDGTLCLVSGSRTFFGC 555
Db 636 PYPEPDDAQKCELCQSADTGDVVFNNQVVDHGRCSYRDP-----YVSCARGECVPVGC 690
QY 556 DGRMDSQVWDRQVCGGDNSTCSPKSGTAGAREVYVTLVTP-----NLTSVY 607
Db 691 DKEVSGMKADDKCGVCGGDNCHRTVKG--TLGKASKQAQAKLVQIPAGARHIQLEALE 748
QY 608 IANHRPLFTHLAVRIGRYVYVAGQMSISPNTPYSLLEDGRVYRVALTEDRLPLREIR 667
Db 749 KSPHRIVVKN---QVTGSFILNPKGEATRTFTAM----GLEWEDAVEDAK-----ESLK 797
QY 668 IMGPLEQDADIQVRRYRGEYGNLTRPDITFY-----FQPKPQAWY- 710
Db 798 TSGPLPE--AJAILALPPTGG--PRSSLAYKYVHEDLLPLIGSNVLEEMDTVEWAL 853
QY 711 --WAAVRGPCSVSGAGLRWVNTSCLDQAKELVETVQCQSQSQPPAPWEACVLBPC-PP 767
Db 854 KSWA----PCSKACGGGIQFTKYGRRRRDHMVQRHLCDHKRKPRIARRCNQHPCSQP 909
QY 768 YNAVDFGPCSASCGG-GLREPRVRCVDAQSGLLKLTPPARCRAGAQQPAVALETCPQP 826
Db 910 VWYTEWGACSRSCGKLGVTQRTGICLLPLSNGTHKVPKAKACAG-----954
QY 827 CPARWEVSEPSCTSAGGAGLALENETCPGADGLEAPVTEGVSVDKLPAPPCVGMGS 886
Db 955 -----DRPEARRPCLRVP 967
QY 887 CPPGWGHLDATSAGEKAPSPWGSIRTAQAHAHVWTPAAGSCSVSCGRGLMEIRFLCMDSA 946
Db 968 CPAQW-----RLGA-----WS-----QCSATCGEGIQQRVVCRTNA 999
QY 947 LRVFVQBELCGLASKPGSRREVQAVPCPARWQYKLAACSVSCRGVVRILYCARAHGE 1006
Db 1000 -----NSLGHCEGDRPDTVQVCSLPA-----CG-----GN 1024
QY 1007 DGEELILLDTQCQGLPEPEQACSLPCPPRWKVMISLGPSCSACGLGTARRSVACVQLD 1066
Db 1025 HONSTVRADV--WELGTPEGQWVPQSEPLHPIKISSTPC-----TGDRSVFC--- 1071
QY 1067 QGQDVEVDEAAACAAALVRPEASVPCLIADCTYRHHVGTWMEGVSVCQDGIQRRDTCGLGQ 1126
Db 1072 ---QMEVLDRYC-----SIP-----GYHRLCCVSC---IKKASGNPFGPD 1105
QY 1127 AQAPVPADFCQHLPKPVTVRGWAGPCVGQGTPLSLVPHEEAAAPORTTATPAGACQRHL 1186
Db 1106 ---PGPTS---LP-PFSTPG---SPLPGQDPA---DAAEPPGK---PTGSEDHQHG 1146
QY 1187 EPT---GTIDMRGP 1198
Db 1147 RATQLFGALDTSFPG 1161
RESULT 11
AT22_BOVIN STANDARD; PRT; 1205 AA.
ID_ATS2_BOVIN
AC_P79331;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
```

15-WAR-2004 (Rel. 43, Last annotation update)
ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase
with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2) (Procollagen I/II
amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC
I-NP) (Procollagen N-endorpeptidase) (pNPI).
ADAMTS2 OR NPI.
Bos taurus (Bovine).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Oxobiidae; Bovinae; Bos.
NCBI_TaxID=9913;
[1]
SEQUENCE FROM N.A.
TISSUE=Skin;
MEDLINE=97225960; PubMed=9122202;
Collage A., Li S.W., Sieron A.L., Nussgens B.V., Prockop D.J.,
Lapiere C.M.; and expression of bovine procollagen I N-proteinase: a
new member of the superfamily of zinc-metalloproteinases with binding
sites for cells and other matrix components.";
Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
[2]
PARTIAL SEQUENCE.
MEDLINE=95348096; PubMed=7622493;
Collage A., Beschin A., Samyn B., Goebels Y., Van Beeumen J.,
Nussgens B.V., Lapiere C.M.;
"Characterization and partial amino acid sequencing of a 107-kDa
procollagen I N-proteinase purified by affinity chromatography on
immobilized type XIV collagen.";
J. Biol. Chem. 270:16724-16730(1995).
-!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
to fibril assembly. Does not act on type III collagen. May also
play a role in development that is independent of its role in
collagen biosynthesis.
-!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
alpha-1(I) at Pro-Gln and of alpha-1(II) and alpha-2(I) chains
at Ala-Gln.
-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
-!- SUBUNIT: May belong to a multimeric complex. Binds specifically to
collagen type XIV.
-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
-!- TISSUE SPECIFICITY: Enzymatic activity is detected at high level
in all type I collagen-rich tissues such as skin, bones, tendons
and aorta and at low level in brain and thymus. The mRNA levels
were disproportionately high in heart, liver, retina and muscle.
-!- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix.
-!- PTM: The N-terminus is blocked.
-!- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
-!- DISEASE: Defects in ADAMTS2 are the cause of dermatosparaxis, a
recessively inherited disorder characterized by severe skin
fragility and biochemically by the presence in skin of procollagen
incompletely processed at the N-terminus.
-!- SIMILARITY: Belongs to peptidase family M12b.
-!- SIMILARITY: Contains 1 disintegrin-like domain.
-!- SIMILARITY: Contains 1 PLAC domain.
-!- SIMILARITY: Contains 4 TSP type-1 domains.
-!- CAUTION: Has sometimes been referred to as ADAMTS3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
or send an email to license@isb-sib.ch).

EMBL: X96389; CAA65253.1; -.
PIR: T18517; T18517.
MEROPS: M12.301; -.
InterPro: IPR001762; Disintegrin.

DR InterPro: IPR006025; Pept_M_Zn_BS.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR002870; Peptidase_M12B_N.
DR InterPro: IPR000884; TSP1.
DR Pfam: PF01562; Pep_M12B_propep; 1.
DR Pfam: PF01421; Reprolysin; 1.
DR Pfam: PF00900; tse_1; 4.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0900; PLAC; 1.
DR PROSITE; PS0092; TSP1; 4.
DR PROSITE; PS0142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Collagen degradation; Extracellular matrix.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 253 BY SIMILARITY.
FT CHAIN 254 1205 ADAMTS-2.
FT DOMAIN 254 473 METALLOPROTEASE.
FT DOMAIN 474 554 DISINTEGRIN-LIKE.
FT DOMAIN 555 610 TSP TYPE-1 1.
FT DOMAIN 612 716 CYS-RICH.
FT DOMAIN 717 845 SPACER.
FT DOMAIN 848 906 TSP TYPE-1 2.
FT DOMAIN 908 968 TSP TYPE-1 3.
FT DOMAIN 969 1023 TSP TYPE-1 4.
FT DOMAIN 1053 1091 PLAC.
FT DOMAIN 31 35 POLY-ALA.
FT DOMAIN 177 180 ZINC (CATALYTIC) (POTENTIAL).
FT METAL 402 402 BY SIMILARITY.
FT ACT_SITE 403 403 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT SITE 685 687 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1092 1092 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 1144 1144 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1205 AA; 133887 MW; 7B5B232A45320371 CRC64;
Query Match 12.3%; Score 955; DB 1; Length 1205;
Best Local Similarity 27.6%; Pred. No. 5.7e-51;
Matches 297; Conservative 130; Mismatches 378; Indels 272; Gaps 46;
QY 91 QROQRRAAGGILHLELLVAVGPDVFOAH-QEDTERVYLTNINIGAEILLRDPISLGAQFRV 149
DB 248 RRRMRHAADDYNNIEVLLGVDDSVQVFGHTEHYVKYLLTNVINEIYHDESIGHINV 307
QY 150 HLKVKVILTEPEGAPNI-TANLTSSLLSVCCWSQ-TINPEDDTDPGHADLVLYTRFDLE 207
DB 308 VLVRILLTSYKMSLSLEIGNPSOLENCWAVLQKRPDTHDEYH-DHAFLTRQDF- 365
QY 208 LPDENRQVRGVTOLGACSPWCSLTIEDTGFDLGVTTIAHEIGHSGFLEHGDGAPSGCGP 267
DB 366 ---GFSGQGVAPVTGMCHPVRSLNHEDGFSSAFVVAHETGHVGLGMEHDG-QGNRCGD 421
QY 268 S---GHVNSDGAAPRAGLAWSPCRRQLLSLAGRARCWDPFRPQPSAGHPDPAQ- 323
DB 422 EVRLGSIAPLVQAFAHFRHWSRCSQBELSYLHS-----YDCLURDDPFTHDWPAIPQL 475
QY 324 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSLLVPLLDGTEC 383
DB 476 PGLYSANEQCRDFGLGYMNMCTAFRT-FDPCQQLWC-SHPDNPYFCKTKGPPDGTWC 533
QY 384 GVEKWCCKGRCSRSLVELTP-IAAVHGRWSWGPRSPCRSCGGVYVTRRRQNNPRPAPFG 442
DB 534 APGHKCFKHC---IWLTPDILKDGNGWAGNSPFGSCSRTCTGTVKFRTRQCDNHPFANG 590

QY	443	GRACVGRADLAQEMCNVQACEKTOLEFMQCOQARTGQQLPSSPGGASFYHGAAPVHSGQ	502
Db	591	GRTCSGLAYDFOLCNSQDCPDALAFREQCQWD---LYFHHGDAQ-HHW---LPHEHR	643
QY	503	DA--LCRHM--RAIGESFIMKRGDSFLDGTFCMFSGPREDTLSLCVSGSCTRFFCQDGR	558
Db	644	DAKERCHLYCESKETGEVYSMKR--MWHGTRC-----SYKDAFSLCVRGDCRKGVCQDGV	696
QY	559	MDSQVWDRQCVQCGDNDSTCSPRKGSFT-AGSAREVVTFLTPTNLTSVVIANRPLPFT	617
Db	697	IGSKQEDKCGVCGGDHSHCKVKGTFRSRPSKLGIIKMFEPAGARHLLIQEADTTSHH	756
QY	618	LAVR--IGGRYVAGKMSISNT-TYPSL-----LEDGRVEYRVALTEDRLPLEER	667
Db	757	LAVKNETGTRKFIENEDVDNSKTFIAMGVWEYRDEDDR-----ETLQ	801
QY	668	INPGL-----QEDADIQVRY-----GEEYGNLTRPDITFTFYFQPKPQAV	710
Db	802	TMGFLHGTITVLVPEGDARISLTKYMHEDSLNVDDNLEDDSV-----GYE	851
QY	711	WAAVR-GPCSVSCGAGRWVNSCLDQARKELVEIVQCGSQOPPAWPAACVLEPC-PPY	768
Db	852	WALKWSPCKSPCKGGGQFTYGCRRRLDKNVHRGFCDSVSKPAIRTCNPQECSPV	911
QY	769	WAVGDFPCGSCASCG-GGLRERPVRCVE-AQGSLLKTLPPARCAGAAQPAVALETCNP	826
Db	912	WVTGEWEPSCSRTGQVRSVRCVQPLHNTTRSVHTKCHNDARPE---GRRACNREL	968
QY	827	CPARWEVSPSCSAGGAGLALENETCPVAGDGLEAPVTEGSGVDEKLPAPPCVGM	886
Db	969	CPGRW-----RAGS-----WS-----QCSVTCGNGTQERPVLCKRTAD	1000
QY	887	CPGWGHLDATSAGEKAPSPMGSTRTGAQAAHVWTPAAGSCSVSCGRGLMELRFLC--WD	944
Db	974	-----RAGS-----WS-----QCSVTCGNGTQERPVLCKRTAD	1000
QY	945	SALRVPVQEEELCGLASKPGSRRVCOAVPCPARWQYKLAACSVSCGRGVRRILYCARAH	1004
Db	1001	DSFGVCREE-----RPETAR-ICRLGPCP-----RN	1025
QY	1005	GEDDGEIILDTQCGLPREPPEP-----QEACSLPECPFRKVMVSLGPCSASCGLGTARR--	1058
Db	1026	TSPDSKSYV---VQWLSRPDPNSFVQETSSKGRQ-----GDKSVFCRNEVLSRYC	1074
QY	1059	-----SVAC-----VQDQGOQVEVDEAACALVRPEASVPCLIAD	1094
Db	1075	SIPCYNKLCKSCNPHDNLTDVDDRAEPPSGKHNDIE-----LMPTLSVPTLVME	1125
RESULT 12			
ATSI_RAT	STANDARD; PRT; 967 AA.		
ID	ATSI_RAT		
AC	QSWUQ1; Q9ER11;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase		
DE	with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).		
GN	ADAMTS1.		
OS	Rattus norvegicus (Rat).		
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;		
RA	Liu X., Tu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,		
RA	Little S.P.;		
RT	"Induction of a disintegrin and metalloproteinase with the		
RT	thrombospondin type I motif (ADAMTS).";		
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE OF 18-967 FROM N.A.		
RC	STRAIN=Sprague-Dawley; TISSUE=Liver;		
RX	MEDLINE=20304099; PubMed=10847486;		
RA	Diamantis I., Luetthi M., Hoesli M., Reichen J.;		
RT	"Cloning of the rat ADAMTS-1 gene and its down regulation in		
RT	endothelial cells in cirrhotic rats.";		
RL	Liver 20:165-172(2000).		
CC	-1- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be		
CC	involved in its turnover. Has antigenic inhibitor activity (By		
CC	similarity). Active metalloproteinase, which may be associated with		
CC	various inflammatory processes as well as development of cancer		
CC	cachexia. May play a critical role in follicular rupture (By		
CC	similarity).		
CC	-1- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu- -Leu-1684		
CC	site, within the chondroitin sulfate attachment domain.		
CC	-1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular		
CC	matrix (By similarity).		
CC	-1- INDUCTION: Down-regulated in endothelial cells derived from		
CC	cirrhotic liver.		
CC	-1- DOMAIN: The spacer domain and the TSP type-1 domains are important		
CC	for a tight interaction with the extracellular matrix. (By		
CC	-1- PTM: The precursor is cleaved by a furin endopeptidase (By		
CC	similarity).		
CC	-1- SIMILARITY: Belongs to peptidase family M12B.		
CC	-1- SIMILARITY: Contains 1 disintegrin-like domain.		
CC	-1- SIMILARITY: Contains 3 TSP type-1 domains.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL; AF149118; AAD34012.1; --		
DR	EMBL; AF304446; AAG29823.1; --		
DR	MEPROS; M12.222; --		
DR	InterPro; IPR006586; ADAM_cysteine.		
DR	InterPro; IPR001762; Disintegrin.		
DR	InterPro; IPR006025; Pept M Zn BS.		
DR	InterPro; IPR001590; Peptidase_M12B.		
DR	InterPro; IPR002870; Peptidase_M12B_N.		
DR	InterPro; IPR000884; TSPI.		
DR	InterPro; IPR008085; TSP 1.		
DR	Pfam; PF01562; Pep_M12B_propep; 1.		
DR	Pfam; PF01421; Reptolysin; 1.		
DR	Pfam; PF00090; tspl_1; 3.		
DR	PRINTS; PRO1705; TSFIREPEAT.		
DR	SMART; SM00608; ACR; 1.		
DR	SMART; SM00209; TSPI; 3.		
DR	PROSITE; PS0215; ADAM_MEPRO; 1.		
DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.		
DR	PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.		
DR	PROSITE; PS00142; ZINC_PROTEASE; 1.		
DR	PROSITE; PS50092; TSPI; 3.		
KW	Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;		
KW	Repeat; Extracellular matrix; Heparin-binding.		
FT	SIGNAL 1 54		
FT	PROPEP 55 252		
FT	CHAIN 253 967		
FT	DOMAIN 253 475		
FT	DOMAIN 476 558		
FT	DOMAIN 559 614		
FT	DOMAIN 616 724		
FT	DOMAIN 725 857		
FT	DOMAIN 854 910		
FT	DOMAIN 911 967		
FT	DOMAIN 194 198		
FT	DOMAIN 205 205		
FT	SITE 205 205		
FT	METAL 401 401		
FT	ACT_SITE 402 402		

FT	METAL	405	405	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	411	411	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	547	547	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	720	720	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	764	764	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	945	945	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	21	21	I -> V (IN REF. 2).
FT	CONFLICT	26	31	KFRSSQ -> RSRGSL (IN REF. 2).
FT	CONFLICT	49	49	V -> A (IN REF. 2).
FT	CONFLICT	72	72	R -> P (IN REF. 2).
FT	CONFLICT	79	79	L -> TR (IN REF. 2).
FT	CONFLICT	249	249	R -> G (IN REF. 2).
FT	CONFLICT	262	265	TMVL -> NLK (IN REF. 2).
FT	CONFLICT	607	607	S -> F (IN REF. 2).
FT	CONFLICT	936	936	L -> V (IN REF. 2).
FT	CONFLICT	962	962	I -> T (IN REF. 2).
SQ	SEQUENCE	967 AA;	105705 MW;	F93C864F6DCDB4CF CRC64;
Query Match				
Best Local Similarity 28.7%; Pred. No. 1.7e-50;				
Matches 251; Conservative 111; Mismatches 373; Indels 139; Gaps 30;				
QY	4	PTPKAPSHAPLLGLALLRHQRHPRARCPPLCVAGILACGLLGCWGPSHFQSCLOA	63	
Db	177	PAEPKEESIAPP-----RFHILRRRR-----GSGGAKCGVM-----DEETLPT	215	
QY	64	LEPQAVSYLSPGAPLK-----GRPSFGFORQORQRAAGILHLELLVAVGPDPV	116	
Db	216	SNSGRESQNTPDQWPLRNPPTQGAGKTPGQ-----SINKKRFVSPRVETNVLVADQSM	271	
QY	117	QAQOETERYVLNINIGALLRDPSPGAQPRVHLVKMVLTEPBGANITANLTSLLS	176	
Db	272	DFHSGGLKYLTLFVAARPYKPSIRNSLSLVVKILVIYEQKGPBEVTSNAALTUN	331	
QY	177	VCWQSOTINPEDDTPGHADLVLYITRFDLPDGNRQVRGTQVGAGCSPTWSLITIED	236	
Db	332	FCSQWQKNSPSDRDPHYDTAIFTQDL-CGSHCTDLGWADVTGCDPERSCSVIED	390	
QY	237	TGFDLGVITAHETCHSFLGHDGAPG-----SGCGPSGHVMAS-----DGAAPRAGLWSP	287	
Db	391	DGLQAFTTAHELGHVFMHPDDAKHACASFNQSGDLSHMASLMSLSDHSQP-----WSP	445	
QY	288	CSRRLLSLSAGBARCVWDPDRPPQPGSAGHP-----PDAQPLIYSANBQCRVAFGP	344	
Db	446	CSAYMTSFLDNGHGECLMDKPQ-----NPKLPDLPTGLYDANRQCTQTFGEESTH	498	
QY	345	CTFAREHLDWCQALSCHTDPLDQSCSRLLVPLLDGTGCGVEKWCCKGRCLVELTPIA	404	
Db	499	CPDA-----ASTCSTLWCTGSGLLVCQTKHFPWADTSCGEGKWCVSKVNTDKHFA	555	
QY	405	A-VHGRWSSGPRSPCSRCGGVVTTRRRQCNPRPAPFGACVAGADLAQEMCNTQAC-E	462	
Db	556	TPVHSGWGPNGPWGDCSRTCGGVQVYTWRECDNPVPKNGKYGCEGKRVYRSCNIEDCPD	615	
QY	463	KTOLEFMSQOQARTDGOFLRSSPGASFYHGAAPVHSGQDLCALCHRMCAIGESFIMKRG	522	
Db	616	NNGKTFREEQC-EAHNEFSAFGNEPTVETPKYAGVSPKDRCKLTCEAKGIGYFPVLQ	674	
QY	523	DSFLDGTFCWPSGPRDGLSLCYSGSCRTGCDGRMDSQOVWDRQVCGGDSNCSPRK	582	
Db	675	PKVVDGTPCSP-----DST-SVCVQGVQCVKAGCDRIIDSKKPKFCGCVCGGNGSTCKKIS	728	
QY	583	GSFTAGRAREXVTELTTPNLTSSVYIANRPLFTH-----LAVRIG-GRYVWAGKMSISP	636	
Db	729	GTVTSTP-GYHDVITTPAGATNIEVKENPRGSRNNGSFLAIRADGTYYILNG-----	781	
QY	637	NTTYPSELDEGRVRYVALTEDRLPRLEIRWGPLOQDADIQVYRRYGEYGNLTRPDI	696	
Db	782	NFTLSTLEQDLYGTVLRYSGSSAALERISFSLKPELTTQVLM-----VGHALRPKI	836	
QY	697	TFIYFQKPRQ-----RWV-----WAAVRGFCVSCAGLRWNYSLDQARLQVET	744	

Db	837	KTYFMKKTEPFPNAIPTSEWIEW-----GBCSKTCGSG--W-----QRRVVEC	881	
QY	745	VQCGQ-----SQPPAWPEACVLEPCPPYAVGDFPCASCGGRLRRPVRVCVEAQG	797	
Db	882	RDINGHPASECAKEVKEPASTPCADLPC-PRWQVDWSPSCSKTCGKGYKRTLKLSHDG	940	
QY	798	SILKTLPPARCAGAAQPAVALTCNPPCPARW	831	
Db	941	GVLSN-----ESCDPLKPKHY	957	
RESULT 13				
ATSI_MOUSE				
ID	ATSI_MOUSE	STANDARD;	PRT;	968 AA.
AC	P97857; C54768;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).			
GN	ADAMTS1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/SVJ.			
RC	MEDLINE=98110583; PubMed=9441751;			
RA	Kuno K., Lizasa H., Ohno S., Matsushima K.;			
RT	"The exon/intron organization and chromosomal mapping of the mouse			
RT	ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";			
RL	Genomics 46:466-471(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	MEDLINE=97150761; PubMed=8995297;			
RA	Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.;			
RA	Matsushima K.;			
RT	"Molecular cloning of a gene encoding a new type of metalloproteinase-			
RT	disintegrin family protein with thrombospondin motifs as an			
RT	inflammation associated gene.";			
RL	J. Biol. Chem. 272:556-562(1997).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Limb, and Mammary gland;			
RC	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,			
RA	Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Vallalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J.J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,			
RA	Schmurch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			
RT	human and mouse cDNA sequences.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[4]			
RP	CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.			
RC	MEDLINE=99303657; PubMed=10373500;			
RA	Kuno K., Terashima Y., Matsushima K.;			
RT	"ADAMTS-1 is an active metalloproteinase associated with the			
RT	extracellular matrix.";			
RL	J. Biol. Chem. 274:18821-18826(1999).			

[5]
FUNCTION.
MEDLINE=20389568; PubMed=10930576;
Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
Ohno H., Matsushima K.,
"ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
FEBS Lett. 478:241-245(2000).
[6]
FUNCTION, AND INDUCTION.
MEDLINE=20243757; PubMed=10781075;
Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
Richards J.S.,
"Progestosterone-regulated genes in the ovulation process: ADAMTS-1 and
cathepsin L proteases.";
Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
-!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
involved in its turnover. Has angiogenic inhibitor activity (By
similarity). Active metalloprotease, which may be associated with
various inflammatory processes as well as development of cancer
cachexia. May play a critical role in follicular rupture (By
similarity).
-!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-|-Leu-1692
site, within the chondroitin sulfate attachment domain.
-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix.
-!- INDUCTION: Induced in vitro in colon adenocarcinoma cells by
interleukin-1, or in vivo in kidney and heart by
lipopolysaccharide. Also induced by LH stimulation in granulosa
cells of preovulatory follicles.
-!- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix.
-!- PTM: The precursor is cleaved by a furin endopeptidase.
-!- SIMILARITY: Belongs to peptidase family M12B.
-!- SIMILARITY: Contains 1 disintegrin-like domain.
-!- SIMILARITY: Contains 3 TSP type-1 domains.
-!- CAUTION: Ref.2 sequence differs from that shown due to a
frameshift in position 7.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.ebi.ac.uk/ebis/sib.ch/announce/>
or send an email to license@ebi-sib.ch).

EMBL; AB001735; BAA24501.1; ALT INIT.
EMBL; D67076; BAA11088.1; ALT_FRAME.
EMBL; BC040382; AAH40382.1; -.
EMBL; BC050834; AAH50834.1; -.
MEROPS; M12.222; -.
MGD; MGI:109249; Adamts1.
InterPro; IPR006586; ADAM_cysteine.
InterPro; IPR001782; Disintegrin.
InterPro; IPR006025; Pept_M_zn_BS.
InterPro; IPR001590; Peptidase_M12B.
InterPro; IPR002870; Peptidase_M12B_N.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP1.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; tsp.1; 3.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00608; ACE; 1.
SMART; SM00209; TSP1; 3.
PROSITE; PS02015; ADAM_MEPPO; 1.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS02014; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS0092; TSP1; 3.
PROSITE; PS00142; ZINC_PROTEASE; 1.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Extracellular matrix; Heparin-binding.

FT	SIGNAL	1	48	POTENTIAL.
FT	PROPEP	49	253	
FT	CHAIN	254	968	ADAMTS-1.
FT	METALLOPROTEASE.			
FT	DOMAIN	254	476	
FT	DOMAIN	477	559	DISINTEGRIN-LIKE.
FT	DOMAIN	560	615	TSP TYPE-1 1.
FT	DOMAIN	618	725	CYS-RICH.
FT	DOMAIN	726	850	SPACER.
FT	DOMAIN	855	911	TSP TYPE-1 2.
FT	DOMAIN	912	968	TSP TYPE-1 3.
FT	DOMAIN	915	199	POLY-ARG.
FT	SITE	206	206	CYSTEINE SWITCH (POTENTIAL).
FT	METAL	402	402	ZINC (CATALYTIC) (BY SIMILARITY).
FT	ACT SITE	403	403	
FT	METAL	406	406	ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	412	412	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	548	548	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	721	721	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	765	765	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	783	783	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	946	946	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	MUTAGEN	403	403	E-Q: LOSS OF ACTIVITY.
FT	CONFLICT	335	335	N -> S (IN REF. 2).
FT	CONFLICT	425	425	T -> S (IN REF. 2).
SQ	SEQUENCE	968 AA;	105841 MW;	42EBDA55499FB6C1 CRC64;

Query Match 12.2%; Score 944.5; DB 1; Length 968;
Best Local Similarity 30.1%; Pred. No. 28-50;
Matches 239; Conservative 103; Mismatches 334; Indels 117; Gaps 27;

QY	81	GRPPSPGFORQORQRAAGILHLELLVAVGPDVFOAHQEDTERVLTNINIGALLRD	140
Db	241	GRPSGEG-----SIRKRFVSSPRVETMLVADQSMADFHGSLGHYLLTLFSAARFYKH	296
QY	141	PSLGAQFRVHLVFWILTEPEGAPNITANLTSSLLSVCGNSQTINPEDDTDFGHADLVLY	200
Db	297	PSIRNISLSVVKILVIYERQKGPETVTSNAALTIRNFCNWKQHNPSDRDEHYDTAIL	356
QY	201	ITRFDLELDGNEQVRGVTOLGACSTNSCLITEDTDFGLGVITIAHIGHSFGLHDCGA	260
Db	357	FTRQDL-CGSHTCDTLGMADVGTCDPSRSCVIEDDGLQAAFTAHGLGHVFNPHDDA	415
QY	261	PG-----SGCGSPSHVNAS-----DGAAPRAGLAWSPCSRROLLSLLSAGARCVMDPPRP	311
Db	416	KHCASLNGVTGDSHLWASMLSSLDHSQP-----WSPCSAYNVTSLDNHGHCELMKDPK	469
QY	312	QPGSAGHP---PDAQGLYTSANEQCRVAFQKAVACTFAREHLDMDCOALSCHTDPLDOS	368
Db	470	-----NPVKLPDDLPGTLYDANRQCOFTGEESKHCPDA---ASTCTTLWCTGSGGLL	520
QY	369	SCSRLLVPLLDGTGTCGVKRCRSLVELTPIAA-VHGRWSSWGSPRSPCSRSCGGGV	427
Db	521	VCQTKHFPWADGTSCEGKCVSKYKNTDMKHFPATPVHGSWGSPGWDGCDSCRTCGGV	580
QY	428	VTRRQCNPRPAFGGRACVADLOAEMCNTOAC-EKTOLEFMSQOCHARTDQGPLRSSPG	486
Db	581	QYTMRECDNPVPKNGKCYCEKRVYRSCNIEDCPDNGKTFREBQC-EAHNEFSKASFG	639
QY	487	GASFYHGAAPVSHSQGDALCEHMCRAIGESFIMKRGDSFLDGTGRCMPSGPREDDTSLCV	546
Db	640	NEPTVEWTPKIAGVSPKDRCKLCEAKGIGTFFVLQPKVVDGTPCSP-----DST-SVCV	693
QY	547	SGSCTFGCDGRMDSQOVMDRCQVCGGDNSCTSPKSGSFTAGRAREYVYVTLVTNLTSV	606
Db	694	QGCQVAGCDRIIDSKKKFKCGVCGGNGSTCKKWSGIVTSTRP-GYHDIIVTIPAGATWI	752
QY	607	YIANH-----RPLFTHLAVRIG-GRYVAVGKMSISPNITYPSLLEDGVEYRVALTEDRL	660
Db	753	EVKRNQGRSGRNGSFLAIRAADGTYLNG-----NFTLSTLEQDLATVKGTVLYRSGSS	806
QY	661	PRLEETIRNGPLOEDADIQVYRRYGEYGNLTRDITFTYFQPKPQA-----WV-710	
Db	807	AALERIRSFSLKEPTLIQVLM-----VGHALRPKIKFTYFMKKKTSTESFNAITPTSEWVI	861


```
Query Match      12.1%; Score 941; DB 1; Length 1211;
Best Local Similarity 29.5%; Pred. No. 4.2e-50;
Matches 276; Conservative 119; Mismatches 390; Indels 150; Gaps 38;

QY 91 QFORQRAAGGILHLLELVAVGPDVFAQH-QEDTERVVLNLTNNTGAELLRDPISGAQPRV 149
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 254 RRRARRHAADDVNIIEVLGVDSDSVGFHGKEHVQKYLTLNINVEIYHDSIGAHNV 313
QY 150 HIAKVVILTEPEGAPNI-TANLTSSLLSVCGWSTINPEDTDFGHADLVLYIRFDLEL 208
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 314 VLVRIILLSYGSMISIEIGNPSQSLNVCRWAYLOQKPDTCGHDEYHDAIFLTRQDF-- 371
QY 209 PDGNEVRVGTOLGGACSPWLSCLITDTCFLGVTIAHEIGHGSHFGLHGDGAPSGCGPS 268
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 372 --GPSMGQYAPVTGCHPVRSTCLNHEDGFSSAFVVAHEHGVHGMHGDG-QGNRCODE 428
QY 269 ---GHVMSDGAAPRAGLAWSPCSRQLLSLSAGRCVWD-----PPRPQGSAGH 318
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 429 VRLGISIMAPLVAFAHFRHFWRSQSQELSRVLSH--YDCLLDDPFAHDWPALPQ----- 480
QY 319 PDAQPLGLYSANEQCRVAFGPKAVACTFAREHLDMQALSCHTDPLDQSSCSRLVPLL 378
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 481 ----LPGLHYMNEQCRDFGLGYMCTAFRT-EDPCQLWC-SHPDNPYFCKTKKGPPL 534
QY 379 DGTGCGVEKWCGRCSLSVELTP-IAAVHGRWSSMGPRSCSFGGVVTRRQCNPP 437
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 535 DGTMCAPGXKCFKHC---IWLTPDLIKRSGWAGSPFGSCRTCGTVKFRTRQCDNP 591
QY 438 RPAFGACVAGDLQAMCMNTQACEKTLQELFMSCQARTDQQLRSSPGGSAFYHWAAGV 497
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 592 HPANGRTCSGLAYDQLCSRQDPSDLAPFREQCQWD---LYFEHGDQ-HW---L 644
QY 498 PHSQGD--LCRMHC--RAJGEFIMKRGDSFLDTRCMPSPREDGTLSCVSGSRTF 553
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 645 PHEHRDAKERCHLYCESRETGEVVMKR--MVHDGTRC-----SYKDAFSLCVRGDCRK 697
QY 554 GCDGRWDSQVWRCQVCGGDNSTCSPKGSFT-AGRAREVYFLTVTPNLTSTVIANHR 612
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 698 GCDGVIGSSQKQDKCGVCGGDNHCKVVGFTTRSPKKGHYKMFEPAGARHLHIEVD 757
QY 613 PLFTHLAVR--IGRYVW-----AGKMSISPTNTPSLELGRVEYRVALTEDRLPR 662
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 758 ATSHHLAVKMLTGKFLINENDVDASKTFIANGVEWYEDEDGR----- 803
QY 663 LBEIRIWGLQDADIQVYRRYGEYGNLTPDITFTYF-----QPKPRQAW 709
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 804 -ETLQTMGLPHGTITVLVI-----PVGD-TRVSTYTKYMIHEDSLNVDNVLDEDSVY 856
QY 710 VMAAVR-GPCSVSCAGLRWVNSCLDQARKELVETVQCQSQPPAWPEACVLEBPC-PP 767
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 857 EWALKWSPSCFKCGGSGQFTKYGCRRLDHWVHRGFCALSKPKAIRACNPOECQOP 916
QY 768 YNAVGFPGCSASCG-GGLERPVRCVE-AQGSLLKTLPPARCAGAAQPAVALETCPQ 825
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 917 VMTGWEPSCQCGRTGMVRSVRCIQLPHDNTTRSHAKHCNDARPE---SRPACSR 973
QY 826 PCPARVEVSPSSCTAGGAGLALENETCVPGAGLEAPVTEGGSVDEKLP-APEPCVG 884
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 974 LCPGRWAGPWSQCVTCGNGTQRPVPCRTADDSF-----GICQERPETARTCL 1025
QY 885 MSCFPGWGHLDATSAGEKA-----PSPWGSIRTGAQAFVMTPAAGSCSVSCGRGIM 936
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1026 GPCP-----RNISDPSKSYVQWLSRPDPSPIRKISSKHG-----COGDKSIFC----- 1071
QY 937 ELRFLWDSALRVPVQBELCGLASKPSRRREVCQA 971
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 1072 -----RMEVLSRYCSI---PGYNKLSCKS 1092

RESULT 15
ATSI HUMAN
ID ATSI HUMAN
AC Q9UH8; Q9NSJ8; Q9P2K0; Q9UH83; Q9UP80;
```

```
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1) (METH-1).
GN ADAMTS1 OR METH1 OR KIAA1346.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
RT "Cloning, characterization and mapping on human chromosome 21 of the
RT orthologue of murine Adams-1.", to the EMBL/GenBank/DBJ databases.
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND FUNCTION.
RA Tissue=Heart;
RX MEDLINE=99367466; PubMed=10438512;
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
RA Lombardo M., Iruela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT family of proteins with angio-inhibitory activity.",
RL J. Biol. Chem. 274:23349-23357(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Tissue=Endothelial cells;
RX MEDLINE=20247184; PubMed=10785405;
RA Glienke J., Schmitt A.O., Pillarsky C., Hinzmann B., Weiss B.,
RA Rosenthal A., Thierauch K.H.;
RT "Differential gene expression by endothelial cells in distinct
RT angiogenic states.",
RL Eur. J. Biochem. 267:2820-2830(2000).
RN [4]
RP SEQUENCE FROM N.A.
RA Tissue=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.",
RL DNA Res. 7:65-73(2000).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA Park H.-S., Toyoda A., Ishii K., Toki Y., Choi D.-K., Groner Y.,
RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
RA Sharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
RA Ramser J., Beck A., Klages S., Hennig S., Rieselmann L., Dagand E.,
RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA Lehrach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.",
RL Nature 405:311-319(2000).
RN [6]
RP SEQUENCE OF 418-967 FROM N.A.
RA Tissue=Melanoma;
RA Blum H., Bauersachs S., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover (By similarity). Has angiogenic inhibitor
CC activity. Active metalloprotease, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1998-Glu-[Leu-1999
CC site, within the chondroitin sulfate attachment domain.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 140.125 Seconds
(without alignments)
3188.394 Million cell updates/sec

Title: US-09-836-712-2
Perfect score: 7751
Sequence: 1 PGRTRPKAPSHSAPLLGLA.....LQSWVPEMQDPQSWKRGKT 1416

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:*

- 1: sp_archea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7600	98.1	1427	4 Q96L37	Q96L37 homo sapien
2	1799	23.2	364	4 Q3UG01	Q9ugq1 homo sapien
3	1185	15.3	223	4 Q9H0G3	Q9h0g3 homo sapien
4	1138.5	14.7	1600	11 Q81IB3	Q81ib3 mus musculus
5	1118	14.4	2165	5 Q19791	Q19791 caenorhabdi
6	1116.5	14.4	1092	11 Q8BKA1	Q8bkai mus musculus
7	1032.5	13.3	1070	11 Q8CG28	Q8cg28 mus musculus
8	1013.5	13.1	1009	11 Q8BK1	Q8bkyl mus musculus
9	933	12.0	759	6 Q8H2M8	Q8h2m8 equus caball
10	907.5	11.7	1059	5 Q9W493	Q9w493 drosophila
11	906	11.7	967	4 Q9NE26	Q9ne26 homo sapien
12	819.5	10.6	900	11 Q8K206	Q8k206 mus musculus
13	759.5	9.8	839	6 Q7YS95	Q7ys95 bos taurus
14	748.5	9.7	833	11 Q8K384	Q8k384 mus musculus
15	748.5	9.7	845	11 Q8BNJ2	Q8bnj2 mus musculus
16	745.5	9.6	1036	11 Q80T21	Q80t21 mus musculus

17	742.5	9.6	1688	5 Q8SXB0	Q8sxb0 drosophila
18	711.5	9.2	623	11 Q8BGP4	Q8bgp4 mus musculus
19	701.5	9.1	769	5 Q8MRL5	Q8mrl5 drosophila
20	687	8.9	1235	4 Q95428	Q95428 homo sapien
21	685	8.8	1280	11 Q9EPX2	Q9epx2 mus musculus
22	674.5	8.7	1572	5 Q44938	Q44938 haenchenchus
23	671.5	8.7	1014	5 Q9V909	Q9v9q9 drosophila
24	671.5	8.7	1353	5 Q9V9Q7	Q9v9q7 drosophila
25	659.5	8.5	1487	5 Q8MPV5	Q8mpv5 caenorhabdi
26	659.5	8.5	1558	5 Q8I7I0	Q8i7i0 caenorhabdi
27	659.5	8.5	2167	5 Q76840	Q76840 caenorhabdi
28	659	8.5	766	4 P82987	P82987 homo sapien
29	657	8.5	1461	5 Q8MYA8	Q8mya8 caenorhabdi
30	614	7.9	3198	5 Q9UG88	Q9ug88 manduca sex
31	599	7.7	2174	5 Q9GQRO	Q9gqro drosophila
32	599	7.7	2772	5 Q9AV4	Q9av4 drosophila
33	599	7.7	2776	5 Q869A0	Q869a0 drosophila
34	599	7.7	2898	5 Q86829	Q86829 drosophila
35	587.5	7.6	1020	5 Q8IU50	Q8iu50 caenorhabdi
36	587.5	7.6	1020	5 Q19204	Q19204 caenorhabdi
37	584	7.5	1043	5 P90884	P90884 caenorhabdi
38	565.5	7.3	957	11 Q7TSK7	Q7tsk7 mus musculus
39	557.5	7.2	951	4 Q86TH1	Q86th1 homo sapien
40	548	7.1	975	5 Q7Z291	Q7z291 caenorhabdi
41	546.5	7.1	951	4 Q60345	Q60345 homo sapien
42	520	6.7	1637	6 Q9XSV8	Q9xsv8 bos taurus
43	520	6.7	5146	6 Q8SPM4	Q8spm4 bos taurus
44	506	6.5	1091	5 Q9W1Z6	Q9w1z6 drosophila
45	493.5	6.4	880	5 Q8MSF8	Q8msf8 drosophila

ALIGNMENTS

RESULT 1

Q96L37 PRELIMINARY; PRT; 1427 AA.

AC Q96L37; 2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13

OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX PubMed=11557746;

RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,

RA Fujikawa K.;

RT "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a

RT Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura.";

RL J. Biol. Chem. 276:41059-41063(2001).

DR EMBL; AY055376; AAL17652.1; -

DR MEROPS; M12.241; -

DR Genew; HGNC:1366; ADAMTS13.

DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.

DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001590; Peptidase M12B.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR00884; TSPL_Zn_BS.

DR Pfam; PF01421; Reprolysin; 1.

DR Pfam; PF00090; tsp_1; 4.

DR SMART; SM00209; TSPL; 7.

DR PROSITE; PS02015; ADAM_MEPPO; 1.

DR PROSITE; PS00092; TSPL_4.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Protease; Signal.

FT SIGNAL; 1 33 POTENTIAL.

```
FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB1BC3AABC1A4442 CRC64;

Query Match
Best Local Similarity 98.1%; Score 7600; DB 4; Length 1427;
Matches 1392; Conservative 0; Mismatches 1; Indels 34; Gaps 1;

QY 24 MHORPRARCPPLVCVAGILACGFLGCGWSPSHFOQSCLOALEPOAVSSYLSGAPLKGKP 83
DB 1 MEQRPRARCPPLVCVAGILACGFLGCGWSPSHFOQSCLOALEPOAVSSYLSGAPLKGKP 60
QY 84 PSPGFORQORQARRAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPSL 143
DB 61 PSPGFORQORQARRAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPSL 120
QY 144 GAQFRVHLVWVILTPPEGAPNITANTSSLLSVCGWSQIINPEDDTPGHADLVLYITR 203
DB 121 GAQFRVHLVWVILTPPEGAPNITANTSSLLSVCGWSQIINPEDDTPGHADLVLYITR 180
QY 204 FOLELPDGNRQVGVYQLGACSPWTSCLITDGFGLGVYTHAIEHIGHSFGLHDGAPGS 263
DB 181 FOLELPDGNRQVGVYQLGACSPWTSCLITDGFGLGVYTHAIEHIGHSFGLHDGAPGS 240
QY 264 GCGPSGHVNASDGAAPRAGLAWPCSRQLLSLSAGRARCVMDPPRPPQPSAGHPDDAQ 323
DB 241 GCGPSGHVNASDGAAPRAGLAWPCSRQLLSLSAGRARCVMDPPRPPQPSAGHPDDAQ 300
QY 324 PGLYYSANQCRVAFQPKAVACTFARHLDMCQALSCHTDPLDOSCSRLVPLLDGTGEC 383
DB 301 PGLYYSANQCRVAFQPKAVACTFARHLDMCQALSCHTDPLDOSCSRLVPLLDGTGEC 360
QY 384 GVEKWKSGRCRSLVELTPTAAVHGRVSSWGPRSPCSRSCGGVYVTRRQCNNRPPAFGG 443
DB 361 GVEKWKSGRCRSLVELTPTAAVHGRVSSWGPRSPCSRSCGGVYVTRRQCNNRPPAFGG 420
QY 444 RACVGADLAQEMCNTQACEKTQLEFMSCQACARTDQQLRSPGSGASFYHMGAAVPHSQGD 503
DB 421 RACVGADLAQEMCNTQACEKTQLEFMSCQACARTDQQLRSPGSGASFYHMGAAVPHSQGD 480
QY 504 ALCRHMCRAIGESFIMKRGDSFLDGTGTCMPSGPRDGTLSLCVSGSCRTFCGDRGMSDQ 563
DB 481 ALCRHMCRAIGESFIMKRGDSFLDGTGTCMPSGPRDGTLSLCVSGSCRTFCGDRGMSDQ 540
QY 564 VMDRCQVCGDNDSTCSRKGSFTAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVRIG 623
DB 541 VMDRCQVCGDNDSTCSRKGSFTAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVRIG 600
QY 624 GRVYVAGKMSISNTTYPSSILEDGRVEYRVALTEDRLPRLEIRTWGPLEQEDADIVYRR 683
DB 601 GRVYVAGKMSISNTTYPSSILEDGRVEYRVALTEDRLPRLEIRTWGPLEQEDADIVYRR 660
QY 684 YGEEYGNLTRPDITFTYFQPKPRAQWMAAVRGPCSVSCGAGLRWVNSCLDQARKELVE 743
DB 661 YGEEYGNLTRPDITFTYFQPKPRAQWMAAVRGPCSVSCGAGLRWVNSCLDQARKELVE 720
QY 744 TVQCQSGQQPPAWPEACVLBPFPYVAVGDFGPCSACSCGGLRERPVRCVEAQSLLKTL 803
DB 721 TVQCQSGQQPPAWPEACVLBPFPYVAVGDFGPCSACSCGGLRERPVRCVEAQSLLKTL 780
QY 804 PPARCAGAOQPAVALETCPNQPAPARWEVSEPSCTSAGGAGLALENETCVFGADGLEA 863
DB 781 PPARCAGAOQPAVALETCPNQPAPARWEVSEPSCTSAGGAGLALENETCVFGADGLEA 840
QY 864 FVTEGPGSVDEKLPAPBPCVGMSCPPGWHGLDATSAGKAPSPWGSIRTAGAAHWVTPA 923
DB 841 FVTEGPGSVDEKLPAPBPCVGMSCPPGWHGLDATSAGKAPSPWGSIRTAGAAHWVTPV 900
QY 924 AGSCSVSCGRLMELRFLCMDALRVPVQEBLCGLASKFGSRREVCQAVPCPARWQYKLA 983
DB 901 AGSCSVSCGRLMELRFLCMDALRVPVQEBLCGLASKFGSRREVCQAVPCPARWQYKLA 960
QY 984 ACSVSCGRGVVRRILYCARAHGEDDGEIILLDTQCQGLPRPBPQACSLPPCPPRWKVMS 1043
```

```
QY 472 QCARDGQPLRSPGASFYHGAAPVHSGDALCRHMCRAIGSFIMKRGDSFLDGTTC 531
Db 121 QCARDGQPLRSPGASFYHGAAPVHSGDALCRHMCRAIGSFIMKRGDSFLDGTTC 180
QY 532 MSPGREDTSLSCVSGSCRTFCDCGRMDSQQVWDRQVCGGDNSTCSPRKGSFTAGRAR 591
Db 181 MSPGREDTSLSCVSGSCRTFCDCGRMDSQQVWDRQVCGGDNSTCSPRKGSFTAGRAR 240
QY 592 EYVTFITVFNLTSVYIANHRPLFTHLAVRIGRYVYVAGKMSISPNTYPSLLEDGRVEY 651
Db 241 EYVTFITVFNLTSVYIANHRPLFTHLAVRIGRYVYVAGKMSISPNTYPSLLEDGRVEY 300
QY 652 RVALTEDRLPRLEIRIINGPQLODADIQV 680
Db 301 RVALTEDRLPRLEIRIINGPQLODADIQV 329

RESULT 3
Q9H0G3 PRELIMINARY; PRT; 223 AA.
ID Q9H0G3
AC Q9H0G3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN DKFZP434C2322.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weil B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ansorge W., Boecker M., Bloecher H., Bauersachs S., Blum H.,
RA Lauber J., Duesterhoeft A., Beyer A., Koehrer K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and
RT Analysis of 500 Novel Complete Protein Coding Human cDNAs.";
RL Genome Res. 11:422-435 (2001).
DR EMBL; AL136809; CAB66743.1; -.
DR InterPro; IPR000859; CUB.
KW Hypothetical protein.
SQ SEQUENCE 223 AA; 24546 MW; 928BFDD77088D28 CRC64;

Query Match 15.3%; Score 1185; DB 4; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.6e-79;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 MRPGQADCAVAGRPLGEVVTURVLESSNCSAGDMLLWGLRTWRQCRKLLDMTFSS 1253
Db 1 MRPGQADCAVAGRPLGEVVTURVLESSNCSAGDMLLWGLRTWRQCRKLLDMTFSS 60
QY 1254 KNTLVVRCRGGGVLLRYGSQLAPETFYRECDMLQFPGWGEIVSPISLSPATSNAGG 1313
Db 61 KNTLVVRCRGGGVLLRYGSQLAPETFYRECDMLQFPGWGEIVSPISLSPATSNAGG 120
QY 1314 CRLFINVAPHARAIHALATNMAGTREGANASYILIRDTSLRTTAFRQGVLYWESESS 1373
Db 121 CRLFINVAPHARAIHALATNMAGTREGANASYILIRDTSLRTTAFRQGVLYWESESS 180
QY 1374 QAEMEFSEGLKAQASLRGQWTLQSVWPEMDPQSNKKGEGT 1416
Db 181 QAEMEFSEGLKAQASLRGQWTLQSVWPEMDPQSNKKGEGT 223

RESULT 4
Q811B3 PRELIMINARY; PRT; 1600 AA.
ID Q811B3
AC Q811B3;
```

```
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloprotease disintegrin 12 protein.
GN ADAMTS-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Cal S., Lopez-Otin C.;
RA "Mouse ADAMTS-12";
RT Submitted (JAN-2003) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AJ537452; CAD60967.1; -.
DR FIR; PT0546; PT0698.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B_N.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000894; TSP1.
DR InterPro; IPR008085; TSP.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR PRINTS; PR00090; tsp.1; 7.
DR SMART; SM00209; TSP1; 8.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Integrin; Protease.
SQ SEQUENCE 1600 AA; 177791 MW; 083089D356E38C5D CRC64;

Query Match 14.7%; Score 1138.5; DB 11; Length 1600;
Best Local Similarity 25.5%; Pred. No. 6e-75;
Matches 336; Conservative 151; Mismatches 503; Indels 327; Gaps 41;

QY 90 RQRQRRAAGGILHLELLVAVGPDVFQAH-QEDTERVVLNLAIGALLRPSLGAQPR 148
Db 237 RRSLSRRSISKRWVETLVADTKTVEYHSENVESYILTMNMVTLGTFHSPSGLNVH 296
QY 149 VHLVKWVILTEPGAPNITANLSSLLSVCSGSTINPEDDTDPGHADLVITRFDLEL 208
Db 297 IVVRLILLEEEOGLKIVHAEKTLSSFCWKQKSNPKSNDLPVPHVDVAVLITKDI-C 355
QY 209 PDGNR--QVRGVTLQGGACSPFTWSCLITEDTGFDLGVTTIAHEIGHSGFLEHDPGSGCG 266
Db 356 AGVNRPCETIGLSQSGMCQPHRSCTNEDSGLPFLAFTIAHELGHSGFIQHDGKE-NDCE 414
QY 267 PSG---HVWSDGAPRAGLANSPCRRQLLSLSAGARCVWDPDPQPGSAGHPDDAQ 323
Db 415 PVGRHPYINSQQIQYDPTPLTWKSKSEYITRFLDRGRGFCDDITPSKKGLKSN---VIA 471
QY 324 PGLYYSANQCRVAFGPKVACVTFAREHLMQALSCHTDPLDQSSCSRLVPLLDGTTC 383
Db 472 PGVIYDVHHQQCQIQYQGNP---TFCQEVENVQCQLWCSV---KGFCRKLDAADGTRC 524
QY 384 GVEKWKSKRCRSLVELTPIAAVHGRWSSWGPSPCRSGCGGVVTRRRQCNRPAPFGG 443
Db 525 GEKKNWAGKCI-TVGKPP-ESIPGGWGRWSPWHSCTCGAGAGASERLCNNPEKFGG 582
QY 444 RACVGADLQAEKMCNTQACEKTLQLEFMSQCCARTDQQLRSPSGGASFYHW----GAAPH 499
Db 583 KYCTGERKRYRLCNVHPCRSDTPTFRQWCSEFTVYKN----QFYRWFPVFNAHP- 636
QY 500 SQGDALCRHWCRAIGSFIMKRGDSFLDGTTCMSPGREDTSLSCVSGSCRTFCGCDGRM 559
Db 637 -----CELYCRPIDEQFSERMLEAVIDGTPCFEGG---NSRNVCIINGICKRVGDDYEI 686
QY 560 DSQVWDRQVCGGDNSTCSPRKGSFTAGRAREYVTFITVFNLTSVYIANHRPLFTHLA 619
```

```
Db 687 DSNATEDRCVCLGDSACQTVKLFQKKEGSGYVDIGLIPKPGARDIRVMEIKAAGFLA 746
Qy 620 VRIIG--RVVAGMISPNVTTPSILEDRVYRVALTRDLRLBEIRIWLQEDAD 677
Db 747 IRSEDEPKYLAGGFTIWNQNYK--LAGTVFQY-----DRKDLKLIAPGPTNESW 798
Qy 678 IQVYRYGEBYGNLTPDITFTYFQK-----PRQAWWAAVR-GPCSVSCGAGLRWV 729
Db 799 LQLLFO-----VTNPGIKYEVTVRKDGLNDVEKLLYFQFGRWTECVTCGTGIRRQ 851
Qy 730 NYSCLDQARKELVETVQCQSQPPAPWPEACVLEPCPPYVAVGDFGPCSASCGG-GLRER 788
Db 852 AHCVKYKH-GIVKTFECNPETQSVQKKCHEKDCPPRWAGWEACSTTCGYGEKR 910
Qy 789 PYRCVBAQSLKLTLPARCRAGAAQPAVALETNPQ-PCPARWEVSPSSCTAGGAGL 847
Db 911 TVLCIQTMSDEQALPATDCQ-HLLKPK-ALVSCNRDILCPSDWTGVNWSVSCGGV 968
Qy 848 ALENETC----- 854
Db 969 RIRSVTCAKNLNEPCDKTRKPNRSLALCGLQCPFSRRVLKPKHDIAPSGKNGQSTAEHDPF 1028
Qy 855 -----VP-----GADGL----- 861
Db 1029 KPIAPTSTRPTLSTPTVPESMSTSTPTINSLGSIASQEDANGWGNNTQABEGSHF 1088
Qy 862 -----EAPVT----- 866
Db 1089 PTSSGSTSQVPTVSWLSIQPDDENVSSAIGPTSEGFWATTTSDSLSSDAMTWQVT 1148
Qy 867 ----- 867
Db 1149 PFYSTWTTDPEVHSGSGSDQPLNKKNSVWIKIGVPEHDPAMETDAELPLGPPP 1208
Qy 882 CVGMSCPPGH-----GHLDATSAEKAPSPGSIKTAQAHHVTPAAGSCSVSCR 933
Db 1209 TSYNGEPPSPPTSKMEGLPAWSFKNETPRDDGM--AEKSRKIPLPLAGDHPHATSE 1266
Qy 934 GLMELRFLCWDLSALR-----VPQELCGLASK-----PGSRREVCQAVPCFARWQV-K 981
Db 1267 KLENHDKLALPNTNPTQGGFVLTEDASNLIAEGLFNLSADYKHLMKDHSPAYWVGN 1326
Qy 982 LAACSVSGRGVVRILYCARAHGEDDGEIILLDTCCQGLPRPEQACSLPFCPPRWKV 1041
Db 1327 WSKCSTTCGLIAYRVSCESSG-----VDADCTTIQPPDPAKKCHLPC-AGWRV 1375
Qy 1042 MSLGPCASCSGLGPARSVACV-QLDQGDYVEDEAACAALVRPEASVPLIADCTYRW 1100
Db 1376 GNWSKCSRNCSGFKIREVQCMDSLDHRSRPPHCQFLAGAPPLSLMSCNLEPCG-EWQ 1434
Qy 1101 VGTWMECSVCGDGIQRRDTCLGPOAQPVPADFCQHLPKPVTVRGC-----WA 1150
Db 1435 VEPWSQCSRSCGGGVQBRGVSC-----PGGLCDWTKRPATTVPNCNRHLCCHWA 1482

RESULT 5
Q19791 PRELIMINARY; PRT; 2165 AA.
AC Q19791; Q27524;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE F25H8.3 protein.
GN F25H8.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_Taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gaiadety S.
RL Submitted (F25-1996) to the EMBL/GenBank/DBJ databases.
```

```
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson M., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell K., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sprout J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL; Z69361; CAAG3288.1; -.
DR EMBL; Z69360; CAAG3287.1; JOINED.
DR EMBL; Z69361; CAAG3287.1; JOINED.
DR PIR; T21371; T21371.
DR HSSP; P15167; 1DTH.
DR MEROPS; M12.135; -.
DR WormPep; F25H8.3; CE05729.
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0008270; Fzinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 14.
DR SMART; SMC0209; TSP1; 18.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 15.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 2165 AA; 24437 MW; FCC3DA8AAA9C4888 CRC64;

Query Match 14.4%; Score 1118; DB 5; Length 2165;
Best Local Similarity 26.6%; Pred. No. 2.9e-73;
Matches 331; Conservative 142; Mismatches 501; Indels 272; Gaps 42;

Qy 91 QROQRRAAGILH-LELLVAVGPDVFOAHQEDTERVYLTNINIGALLRDSLGAQFRV 149
Db 267 RKARSRAANSWHDYVEVLVADTKMVEYHGRSLEDYVLTTFSTVASIYRHOSLRASIN 326
Qy 150 HLVMWILTEPEGAPNTANILTSLLSVCGHSQTINPEDDTPGHADLVLYITRPLELP 209
Db 327 VVVKLIYLVKTENAGPRITQNAQTLQDFCRWQYNDPDDSSVQHHDVAILTRKDCRS 386
Qy 210 DGNRQVEGVTLQGCASPTWSCLITEDTGFGLGVITIAEIGHSGFLEHGDGAPSGGCP-- 267
Db 387 QGKCDTGLAELGTMCMDMKSALIEDNGLGSAFTIAHELGHVFIPIHDE--RKCTYM 444
Qy 268 -----SGHYMSADGAAPRAGLAWSPCSRRLQLSL--SAGRARC 305
Db 445 PVNKVKCFQSTKFKDTQFQNNFHIMAPTELYNTHPWSWSPCSAGMLERLENNRGQTQCL 504
Qy 306 WDPERPQGSAGHPDA-----QGLYISANQCRVAPGKAVACTFAREHLMQALSCH 361
Db 505 FD-----QPVERRYEDVFVRDEPGKKYDAHQCKEYFVGPASELCPY----MPTCRALWCA 556
Qy 362 TDFLDQSSCSRLVPLLDGTGCGVEK--WCSKGRCSLVELTP--IAAVHGRWSSNGPRS 417
Db 557 TFYGSQMGCRTHMPWADGTGTCDESRSMFCHHGAC---VRLAPESUTKIDQGDNRWSG 613
Qy 418 PCSRCGGVVTTRRRQNNRPAFGGRACVAGADLAQEMCNTQACEKTLQEFMSQQCATD 477
Db 614 ECSRTCGGVQKGLRDCDSPRNGGKVCVQGRERYRSNTECPWDTPYREVQCSFEN 673
Qy 478 GQPLRSSPGGASFYHWGAAPVPHSGDALCRHMCRAIGHSEFIMKRGDSFLDGTCTMPSGPR 537
Db 674 NKDIGIQGVASTNTHWPKYANVAFNERCKLYCRILSGSAAFYLLRDKVVDGTGTPCDRNGD- 732
```


QY 538 EDGTLCLVSGSCRTFCGDMDSQVWDRQVCGGDMSTCSPRKGSFTAGRARVYVFL 597
 Db 733 -----DICVAGAMPACGCHLSTLRDKCGVCGGDDSSCKVKGTEQGTGYNVM 787
 QY 598 TYPNLTSVIA-----NHRPLFTHLVR-IGGRYVAGKMSISNTTYPSELLEDRVEY 651
 Db 788 KIPAGSANIDIRKQGYNMKEDDNYLSRAANGFELNGHFQVS-LARQIAFQDTVLEY 846
 QY 652 RVALTREDRLPRLEIRINGFLQEDADICVRRYEEYGNLTRPDITFY-----FQ 702
 Db 847 S-----GSDAI--IERINGTPIRSIDYVHL-----SVGS-HPDTSYEYMTAAVNAVIR 895
 QY 703 P-----WMAARVGPSCVSCAGLR-----KPRQ----- 708
 Db 896 PISSALYLWRVTWTWTECDRACRQOQKMLCLDMSTHRQSHDRNCQVLPKQATRCMN 955
 QY 709 -----WMAARVGPSCVSCAGLR-----KPRQ----- 727
 Db 956 IDCSFTWITEDV-SSCSAKCGSKQRQVSCVKMEGDEQTPASBHLCDRNSKPSDIASCY 1014
 QY 728 -----WNVYS-----CLDQARKELVETVQCGSQOQPPAPWEACV 761
 Db 1015 IDCSGRKNYGEWTSCTSGNSGKMKRKSVCVDSNRRVDESL-C-GREQKEATERECN 1072
 QY 762 LEPCCPYWAGDRGPCSASCGGLRERPRVCVEAAGSLKTLPLPARCRAGAQPVALET 821
 Db 1073 RIPC-PRWYGHWSERSCDGGVYKMRHAQCLDAARETHT---SRC-----GPAQTQEH 1123
 QY 822 CNPQCPARWEVSERSSCSAGAGLALNETCPVGDGLEAPVTEGFGSVDEKLPAPEP 881
 Db 1124 CNEHAC-TWQFGVWSDCSAKGQGVQYRANC---TDRHRSVUPHRCLEKWEKI-ITPK 1178
 QY 882 CVGMSCPFGHLDATSAAGEKAPSPWMSIRTAQAHAHVWTPAAGSCSVSCGRLMELRFL 941
 Db 1179 CHRESCP-----KXKLGE-----WS-----QCSVSCEDGSSRRVS 1209
 QY 942 CMDSALRVPVQBELGLASKPGSRREVCQAVPCPARWQYKLAACSVCGRGVVRILYCA 1001
 Db 1210 CV-SNGTEVDMSLCGTASDRPASHQCNLTGCPFRNTDWSACSVSCGIGHRETTCCI 1268
 QY 1002 RAHGEDDGEIILLDTQCGLPREPPEPQACSLPECPPRKVMSLGPCSASCGLTARRSVA 1061
 Db 1269 YREQSVDA-----SFGCDTKMPETSQTHLLPC-TSMKPSHWSPCSVTCSGIGITRSVS 1321
 QY 1062 CVQLDQGVDEVEAACAALVRPEASVPC-----LIADC-TYRHHVGTWMEC 1107
 Db 1322 CTRGSEG--TIVDEYCDNRTRPLKKTCEKDTCDGPRVLQKQADVPPIWATGFWTAC 1379
 QY 1108 SVSCGSGIQRRTDCLGPOQAQAPVPADFCOHLPKPVTVRGCGWAGPC 1153
 Db 1380 SATCGNGTQRLLKC--RDHVRDLPDFYCNHLDKEVSTNCELRLDC 1423

RESULT 6

Q8BKAL PRELIMINARY; PRT: 1092 AA.
 AC Q8BKAL;
 DT 01-MAR-2003 (T-EMBLrel. 23, Created)
 DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 DE Hypothetical metalloprotease.
 GN El30314N14RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Eye;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 DR EMBL; AK053851; BAC35556.1; --
 DR MGD; MGI:2442600; El30314N14RIK.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP_1.
 DR Pfam; PF01421; Reprlysain; 1.
 DR Pfam; PF00090; tsp_1; 5.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSP1; 5.
 DR PROSITE; PS00215; ADAM_MEPRO; 1.
 DR PROSITE; PS00092; TSP1; 4.
 KW Hypothetical protein.
 SQ SEQUENCE 1092 AA; 120324 MW; B5F03956553536AE CRC64;

Query Match 14.4%; Score 1116.5; DB 11; Length 1092;
 Best Local Similarity 28.7%; Pred. No. 1.5e-73;
 Matches 292; Conservative 129; Mismatches 362; Indels 235; Gaps 34;
 QY 94 QBRRAAG--GILHLELLVAGPDVFOAH-QEDTERVYVLTNINIGAEILLRDSLGAPRV 149
 Db 109 RPRRSAGKSONGLNVELVWADAKWKEKHDDVTYILTVMNVSSLFKDTIGSDINI 168
 QY 150 HLYKVMIL-TEPGAENITANLTSSLSVCGWSQNTINPDTPDGHADLVLTITFDL-E 207
 Db 169 VVVSLLILEEPEGL-LINHHADQSLNSFCQMSALVGNKGRHDAH---ILLTGFIDCS 224
 QY 208 LPGNRQVRVTOGGACSTWCLITDFTDGLVTTAIEHGFHLEHGDGAPSGCGP 267
 Db 225 WKNEPCDTLGFAPISGCMKYSKYSTINEDTGLGLAFTIAHESGHNFVMVHDG-EGNPCR 283
 QY 268 S-CHVWASDGAAPRAGLWPCSRROLLSLSAGRCVWDPVPPRQPSAGHPDAPGL 326
 Db 284 AEGNIMSPITGNNGVFSWSSCSQYLKFKLSTPQAGCLVDEPK-QTGOYKY-PDKLPQG 341
 QY 327 YGSANQCRVAFQKAVACTFAREHLDWCOALSCHTDPLDQSSSRLVPLDGTGECVE 386
 Db 342 IYDADWQKQVQAKAKLCSLG-VKMDICKLWCHR---VGHRCETKMPAAEGTACGLS 397
 QY 387 KWSKGRCSRLVELTPIAAVHGRWSSWGPSPCSRSCGGVVRTRRQCNRPAPFGRAC 446
 Db 398 MWCRCQCCVKLGLGP-RPIHGQWSAMSKWSECSTCGGVKFKQERHCSNPKPQYGGKYC 456
 QY 447 VGADLAQEMCNTOACEKTQLEFMSQCCARTDQPLRSPGSGASFYHMGAAVPHSQGDALC 506
 Db 457 PGSSRIYKLCNINPCPENSLDFRAQCAEYNNKQPR-----GWLRYKPYTKVEEDR-C 510
 QY 507 RHWCRAGIESFIMKRGDSFLDGTFCMPSPGREDGTLSLCVSGSCSTFCGDMDSQVWD 566
 Db 511 KLYCKAENFEFFAMSGVKVDGTPCSPH--RND-----VCIDGICELVGDHGLSKAVSD 564
 QY 567 RCOVCGDNSTCSPKGSF-TAGRAREVVTPLTVPNLTSVVIANHRPLFTHLAVR-IGG 624
 Db 565 ACQVCKGDNSTCKFYKGLYSQHKANEYYPVVTIPAGARSEIQLQLSSSVLAVRSLQ 624
 QY 625 RYVVAGKMSIS-----PNTTYPSELLEDGVEYRVALTEDRLPRLEIRINGFLQEDAD 677
 Db 625 KYLTGWSIDWPDGDTFFAGTTF-----EYORSFN-----RPERLYAAGPTWETLV 670
 QY 678 IQVRRYGEYGNLTR---PDITTYFQPK-----PRQAWTAAVGRPCSCVCGAG 725
 Db 671 FEI-----LTQGNPGIAWKALPKVMNVTPQATKRYHHHTWRTVQSDCVTCGGG 720
 QY 726 LRWVNYSLDQARKELVETVQCGSQOQPPAPWEACVLEPCPPYWAGVDFGPCSACGGGL 785
 Db 721 YISIKALCLRDQHTQ-VNSGFCSVTRKATEPKICNAFSCPAYWLPGEWSACKSCAGGQ 779
 QY 786 RERPVRCVEAQSLLKTLPLPARCRAGAQPVALETCPQPARWEVSEPSCTSAGGA 845

```
Db 780 QSRKIRCVQ-----KKPFQKEEAVLHSLCP-----VSTPTQV----- 811
QY 846 GLALENETCVPGADGLEAPVTEGPGSVDEKLPAPBPCVGMSCPPGWHLDATSGAKEKAPS 905
Db 812 -----QVNCNHACPPPEW-----SPS 826
QY 906 PWGSIIRTAQAHHVWTPAAGSCSVSCGRGLMELRFLCWDLSALRVPVQBELCGLASKPGSR 965
Db 827 PW----- 828
QY 966 REVQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGDDGEBEILLDTQCQGLPRPE 1025
Db 829 -----SQCKTGRGVRRREVLC-----KSPRAETLPESLCSSSPRPE 866
QY 1026 PQACSLPEPCPP-----RWKVMSLGPCSASCGLGATARRSVACVQLD-QGQDVEVDAAC 1078
Db 867 AQEGCVLGRCPKNNRLQWIASAWSECATCGLGVRKRELKCKVEKTLQKLTFFPERRC 924

RESULT 7
Q8CG28 PRELIMINARY; PRT; 1070 AA.
AC Q8CG28;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc metalloendopeptidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Brathwaite M., Nagaraja R., Abe K.;
RT "Genomic Sequence Analysis in the Mouse t-complex Region.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF528163; AA017380.1;
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F-zinc ion binding; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR002870; Peptidase M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprylisin; 1.
DR Pfam; PF00090; tsp.1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SMO0209; TSP1; 5.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 5.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1070 AA; 117203 MW; 6B996425EBA2B8D8 CRC64;

Query Match 13.3%; Score 1032.5; DB 11; Length 1070;
Best Local Similarity 26.4%; Pred. No. 2.4e-67;
Matches 316; Conservative 111; Mismatches 431; Indels 337; Gaps 41;

QY 26 QRPRARCPPLCVAGILACGFLGCGPSPHFQSCGLQALE----- 65
Db 81 ORAARHC-----LYAGHLQOAGSHVAVSTCGGLHGLIIVADDEYLLIEPLQGGPK 132
QY 66 -----POAV-----SSYLSF-----GAPLKGPR-----PSP----- 86
Db 133 GHRCPESGPHVYKRSILRPHLDACVGRDEKPKWGPWWLRTLKPPPARPLGNESR 192
QY 87 ---GFQRQRQRRAAGIGLHLELIVAGPDVFOAH-QEDTERYVLTNLINIGAEILRDFS 142
Db 193 GQLGLKRSVSRR-----YVETLIVADKMVAVHGRDRVQYVLAINNIVAKLFQDSS 245
QY 143 LGAQFRVHLVKQVILTEPEGAPNITANLTSSLLSVCGWSQTI-----NPEDDTPDGH 195
```

```
Db 246 LGNIVNLTRLILLTDDQPTLEITHHAGSLDSFCWKQKSIHSGHGAIPENGVAHH 305
QY 196 DLVLYITRDL-----ELPDGNQRVQVTLGGACSPTWSCLTITGTDGLGVTHAHEICH 251
Db 306 DTAVLITRYDICTYKXKPCG---TLGLAPVGMCEERERSINEDIGLATFTIAHEIGH 362
QY 252 SFGLHDGAPGSCGSPGH-----VMSDGAAPRAGLAWSPCSRRLQLLSLSAGRACVMD 307
Db 363 TFGMNHGCV-GNCGARGQDPAPKLMAAHITKNTNPFVWSSCSRDIYITFSLDGLGLCLNN 421
QY 308 PRPQPGSAGHPDAPGLYISANEQCRVAFGPKAVACTFAREHLDLMQCALSCHTDPLDQ 367
Db 422 RP---PRQPFVYTVAPGQAYDADEQCRFQGVKSRQCKYG---EVCSELWC-----LSK 470
QY 368 SS-CSBLLVPLLDGTEC---GVEX-WCKSGKCRSLVELTPIAA-----VHGRWSSWGPRSP 418
Db 471 SNRCITNSIPAEGTLCQHTIDKWCYKVC-----VPFGSRPEGVDGAWGPTWGD 524
QY 419 CSRS CGGVVTRRRQCNPRPAFGGRACVGDADLAEMCNTQACEKTQLEFMSQCCARTDG 478
Db 525 CSRS CGGVSSSRHCHDSEPTTGGKYCLGERRHRSCTNDCCPGSQDFREMOCSEFDS 584
QY 479 QPLRSSPGGASFYHWGAAPHSQGDALCRHMCRAIGESFIMKRGDSFDLGTCTMSPSGPRE 538
Db 585 VPFR-----GKFTW---KTVRGGGVKACSLTCLAEGFNYTERAAAVVDGTECRPD--- 633
QY 539 DGTLSLCVSGSCRTFGCDGRMDSQQVWDRCQVCGDNSTCSFKSGFT-AGRAREYVTF 597
Db 634 --TVDICVSGECKHVGDRLREDKCRVCGDGSACETIEGVSPALPGTGYEDVV 691
QY 598 TVTPNLTSVYIANHRPLFTHLAVRIGRVYVAGKMSISPNTYPSLLBDDGRVEYRVALTE 657
Db 692 WIPKGSVHIFQDNLNLSHLAKDQESLLEGL---PGTPQPHRLPLAGTTFHLRQGP 748
QY 658 DRLPRLBEIRWGLQEDALIOVRRYGEENLTPDITTYFPQPKRQA-----WVVA 712
Db 749 DQAQSLAL---GFINASLIIMV-----LAQELPALHYRFNAPIARDALPPYSWHVA 798
QY 713 AVRGPCSVSCAGLRWVNYSCLDQARKVELVTCQSQQPPAPWPEACVLEPCPPYWAVG 772
Db 799 PWT-KCSAQACAGSQVQVVECRNQLDSSAVAPHYCSGHSKLPKRQACNTEPCPPDWVVG 857
QY 773 DFGPCASCGGLRERVRVCRVCAQGLSLKTLFPARCRAGAQQPAVALETNCPQCPARWE 832
Db 858 NWSRCSRSCDAGVRSRVVC-----QRRVSAABEKKALDDSDAC-PQPRP--- 899
QY 833 VSEPPSCTSGAGAGLALENETCVPGADGLEAPVTEGPGSVDEKLPAPBPCVGMSCPPG 892
Db 900 -----EVLEACQCPMCPPEWA 915
QY 893 HLDATSAKEXAPSPWGSIRTAQAHHVWTPAAGSCSVSCGRGLMELRFLCMLSALRVPVQ 952
Db 916 TLD----- 918
QY 953 EELCGLASKPGSRREVQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEI 1012
Db 919 -----W-----SECTSPCGPLHRVVLCKSA-----DQST 945
QY 1013 LLDTCQGLPRPEQACSLPEPCPP-RWKVMSLGPCSASCGLGATARRSVACVQLDQGDV 1071
Db 946 LPFGHCLPLPAKPPPTMRCNLRRCPARVWVTEWGECSCTCGGLGQQQRTVCTS-HGQ-- 1002
QY 1072 EVDEAACAALVRPEASVPCLIADCTYRWHVGTWMECSVSC-----GDGIQRDR 1120
Db 1003 --PSRECTEALRP-----STMQOCEAKCDSVPPGDGPECKD 1038

RESULT 8
Q8BKY1 PRELIMINARY; PRT; 1009 AA.
AC Q8BKY1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
```

01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
 ADAMTS-12 precursor (Fragment).
 GN AI605170.
 OS Mus musculus (Mouse).
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=C57BL/6J; TISSUE=Head;
 RC MEDLINE=22354683; PubMed=12466851;
 RX the FANTOM Consortium, RA
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK048612; BAC33391.1; .
 DR PIR; PT0546; PT0698.
 DR MGD; MGI:2146046; AI605170.
 DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
 DR GO; GO:0008270; F-zinc ion binding; IEA.
 DR GO; GO:0006508; P-proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP 1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; tsp 1; 4.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS50215; ADAM_MEPRO; 1.
 DR PROSITE; PS50092; TSP1; 3.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 FT NON TER 1009 1009
 SQ SEQUENCE 1009 AA; 113352 MW; 7B63218CFE0FDB1 CRC64;
 Query Match 13.18; Score 1013.5; DB 11; Length 1009;
 Best Local Similarity 31.8%; Pred. No. 5.6e-66;
 Matches 256; Conservative 114; Mismatches 361; Indels 75; Gaps 27;
 QY 90 RQRRORRAAGGILHLELLVAVGPDVFAH-QEDTERVYLTNLNIGAEILLRDLPSLGAQFR 148
 DB 237 RRSLSRSISKEWVETLVADTKVYHGSSEVSEYILTNMVTGLFHSISGNLVH 296
 QY 149 VHLVVKVILTEPGAPNITANLSSLSVCGWSQTINPEDTDPGHADLVLYTRFDLEL 208
 DB 297 IVVRLILLEEEOGLKIVHAEKTLSSFCWKQKSNPKSGLNVEHDAVAVLITRKDI-C 355
 QY 209 PDGNR-QVETGVLGACACFTWCLITEDGPDGLVGTIAHEIGHSFGLHDGAPSGCG 266
 DB 356 AGVNRPCETLGLSLQMGCPHRSCTNEDSGLEFLAFTIAHELHSGFIQHDGKE-NDCE 414
 QY 267 PSG---HVMAADGAAPRAGLAWSPCSRQLLSAGRARCVMDPPRPPQSGSAGHPDDAQ 323
 DB 415 PVGHPVIMSQIQDYDTPLTWKSKSEYTRFLDRGRGCLLDIPKSKGLKSN---VIA 471
 QY 324 PGLYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLVPLLDGTEC 383
 DB 472 PGVIYDVHQCQLQYGENA---TFQEVENVQTLWCNV---KGFCSKLDAAADGTRC 524
 QY 384 GVEKWSKGRCSRSLVELTPTIAVHGRVSSWGPRSPCSRSCGGVTVRRRCQNPAPAGG 443
 DB 525 GEKKWCWAGKCIIT-VGKKP-ESIPGGRVSPWHSCHSRTCCAGAAQSAERLCNNPPEKFG 582
 QY 444 RACVGADLQAEKMTQACEKTLQELFMSQOQARTQQLRSPGSGASFYHW---GAAVPH 499
 DB 583 KYCTGGERKRVLCNVHPCRSSTPTFRQMCSEFTVTPYKN-----QYRWFFVNAALHP- 636
 QY 500 SQGDALCRHMCRAICGSEFPMKRGDSFLDGTGCMPSGPGREDGTLSLCVSGSGSCRTGCGDGRM 559

QY	748	----	QSGQQPAPWPEACVLEPCPPYWAIVGDFGSCSACGGGLBERPVRVCVEAQSSLKTLTP	804	
Db	971	WTHAKNKRPARSGCGDQCPFAHWPWGQWQCFVTC-----RPYGFV-----	AP	1015	
QY	805	BARCRA----	GAQQPAVALETNCNPQPCPARWEVSEPS	837	
Db	1016	PQRRRSVCLDEHDVVVADAECHLQKPAEMEPCESS	1052		
RESULT 11					
Q8NE26 PRELIMINARY; PRT; 967 AA.					
ID	Q8NE26				
AC	Q8NE26;				
DT	01-OCT-2002	(TrEMBLrel. 22, Created)			
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)			
DT	01-OCT-2003	(TrEMBLrel. 25, Last annotation update)			
DE	Hypochemical protein.				
OS	Homo sapiens (Human)				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
NCBI_TaxID=9606;					
[1]	SEQUENCE FROM N.A.				
RP	TISSUE=Testis;				
RC	Straussberg R.;				
RL	Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.				
EMBL;	BC036515.1; -				
GO;	GO:0004222; F-metalloendopeptidase activity; IEA.				
GO;	GO:0008270; F-zinc ion binding; IEA.				
GO;	GO:0005118; P-selectin transport; IEA.				
GO;	GO:0006508; P-proteolysis and peptidolysis; IEA.				
InterPro;	IPR006586; ADAM cysteine				
InterPro;	IPR001128; Cytochrome P450.				
InterPro;	IPR001590; Peptidase M12B.				
InterPro;	IPR002870; Peptidase_M12B_N.				
InterPro;	IPR006025; Pept_M_Zn_BS.				
InterPro;	IPR000884; TSP1.				
InterPro;	IPR008085; TSP 1.				
Pfam;	PF01562; Pep M12B_Dropep; 1.				
Pfam;	PF01421; Reprolysin; 1.				
Pfam;	PF00090; tsp_1; 3.				
PRINTS;	PRO1705; TSP1REPEAT.				
SMART;	SM00608; ACR; 1.				
SMART;	SM00209; TSP1; 3.				
PROSITE;	PS0215; ADAM_MEPRO; 1.				
PROSITE;	PS0086; CYTOCHROME_P450; 1.				
PROSITE;	PS0092; TSP1; 3.				
PROSITE;	PS00142; ZINC_PROTEASE; 1.				
KW	Hypochemical protein.				
SW	SEQUENCE 967 AA; 105387 MW; FFD399674201C3D CRC64;				
Query Match 11.7%; Score 906; DB 4: Length 967;					
Best Local Similarity 29.8%; Pred. No. 4.7e-58;					
Matches 237; Conservative 111; Mismatches 320; Indels 134; Gaps 31;					
QY	83	PFSPGFORQK-----	QRRRAAGGILHLLIVAVGPDVFOAHQEDTERVLTNINIGA	135	
Db	231	PDQPALQGVQPTGTSGIRKPFVSVHRVYETMLVADQSWAEFHGSLKHYLLTLESVAA	290		
QY	136	ELLRDPISLGAQFRVHLVKMVLITPEGAPNITANLTSSLLSVCGWSQITINPEDDTPGHA	195		
Db	291	RLVKHPSIRNSLVVVVKILVTHDEQKGPETVSNAAALTLRNFQKQHNPPPSDRDAEHY	350		
QY	196	DLVLYITRDELEPDGNR--	QVRGVTLQGCASPWSCLITETGDLGVTHAIEHCHSF	253	
Db	351	DTAILFTRQDL---	CGSQCOTDLGMADVGTCDPSRSCSVIEDDQGAFTTAHELGHVF	407	
QY	254	GLEHDCAGP----	SGCGPSGHVMA-----	DGAAPRAGLAWSPCSRQLLSLSAGRARC	304
Db	408	NMPHDDAKOCASLNGVYNQDSHMMASMLNJDHSQ-----	WSPCSAYMTITSFLDNGHGEC	462	
QY	305	VNDPPRPQFGSAGHP---	PDAQPLGYSIYANEQCRVAFGPKVACTFAREHLDNCQALSCH	361	

```
Db 463 LMDKPQ-----NPIQLPGLDPTGSDYANRQCFTEGDSKHCPSDA---ASTCSTLWCT 512
Qy 362 TPLDSSCSRLVPLLDGTGCGVEKWCGRCHSLVEL-----TPIAAVHGWSSWGRPS 417
Db 513 GTSGGLVCQTKHFPWADGTSCEGKWCINGKCNKTDKHFDPFF---HGSWGMWGPWG 569
Qy 418 PCSRSCGGVVTTRRRCNNPRPAFGGRACVGAQLQAEMCNTOAC-EKTQLEFMSQCCART 476
Db 570 DCSRTCGGVQVYTRBCDNPVKNNGKYGCEGKRVYRSCNLEDGPDNNGKTFREEQC-EA 628
Qy 477 DQQLASSFGGAFYHWGAAPVHQSQDA---LCRHMCRAGIESFIMKRGDSFLDTRCMP 533
Db 629 HNEFSKASFGSGPAVEW---IPKYGVSXPDRCKLICQAKGIGYFVLQPKVVDGTPC-- 683
Qy 534 SGPREGDTLSLVSGSCRTFCGDRMDSQQVWDRCCVCGGDNSTCSPRKGSFTAGRAREY 593
Db 684 ---STDST-SVCVQGCVKAGCDRIIDSKKFKDCGCGGNGSTCKKISGSVTSKAP-GY 738
Qy 594 VIFLTVTNLTYSVIANH---RPLTHLAVRIG-GRYVVGKMSISFNITYPYSLLEDG 647
Db 739 HDIITPTGATNIEVKORNGSRNNGSFLAKAADGYIILNGDYTLG-----TLEQDI 792
Qy 648 RVEYRVALTEDRLPRLEIRIWIQLQDADIQVYRRYGEYCNLTRPDITFTYFQPKPRQ 707
Db 793 MYKGVLRVSGSSAALRISFSPLKPLETLIOVL-----TVGNALRPKIYFYVKKKKE 847
Qy 708 -----AMV---WAAVRGPCVSCGAGLRWVNYSCLDQARKELVETVQCQSQQP-- 753
Db 848 SFNAITFSAWVTEW-----GECSSKSELG-W-----QRELVECRDING--QPAS 890
Qy 754 -----PAWPEACVLEPCPYWAVGDFPGCSACGGGLRBPVYCVQAQSLKLTLP 806
Db 891 ECAKEVKPASTRCPADHPC-FQWQLGEWSSCKTCGKGYKRSKLSLSDHGGVL----- 943
Qy 807 RCEGAGQPAVALETCPQPCP 828
Db 944 -----SHESCDPLKXP 954
```

RESULT 12

```
O8K206 PRELIMINARY; PRT; 900 AA.
ID Q8K206;
AC Q8K206;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ADAMTS16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC034739; AAH34739.1; --
DR MGD; MGI:2429637; Adamts16.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS00215; ADAM_MPRO; 1.
DR PROSITE; PS00092; TSP1; 1.
KW Hypothetical protein.
```

SQ SEQUENCE 900 AA; 100679 MW; 590BE2B0E73740FF CRC64;

Query Match 10.6%; Score 819.5; DB 11; Length 900;
Best Local Similarity 29.6%; Pred. No. 1.1e-51;
Matches 226; Conservative 105; Mismatches 306; Indels 127; Gaps 26;

Qy 3 RPTRPKAPSHSAPLL---GLALLRMHQRHPRAPCPPLCVAGILACGFLLCGCGPSPHFQOS 59
Db 205 RSTERQAPRENEVIMITRKDLARPHLHD-----NPHL---GPSOKHF 246
Qy 60 CLOAL-----EPQAVSSVLSFGAPLKGPPSPGFQFQRRQRRRAAGGILHLELLVAVGPDV 115
Db 247 CGSRKYMPOPNDLLYILFD-----EYKPSRHKSLLKSHRNEELNVELVVDKRM 300
Qy 116 FOAH-OEDTERYVLTNLNIGAEILLRDPGLGAQFVHLVKWVILTEPEGAPNITANLTSSL 174
Db 301 MQSHGHENITYVLTILNVSALFKDGTIGNINIVIVGLILLEDEQPLAISHHADHTL 360
Qy 175 LSYCGHSQITINPEDDTPGHADLV--LYTRFDLELDPGNRQVRQVTOIGGACSPFWSC 232
Db 361 TSFCQWCSGLMGKDGTRHDHAILLTGLDICSWKNEPCD---TLGFAPISGMCSKYRSCT 416
Qy 233 ITEDTGFDLGVTTIAHEIGHSGFLEHGDGAPGSGGSPS-CHVNASDGAAPAGLAWSPCSR 291
Db 417 VNEDSGLGLAFTIAHESGHNFVMDHG-EGNMCKKSEGNIMSPITLAGRNVFSSWSSCSRQ 475
Qy 292 QLLSL-SAGARCVWDPPRQPGSAGHPDPAQGLYYSANEQCRVAFGPKAVAC--TFAR 349
Db 476 YLHKFLSTAQAICLAD--QPKPVKEKYPEKLPGQLYDANTQCKWQFGSKAKLCMLDFRK 533
Qy 350 EHLDMCALSGHTDPLDQSSRLVPLLDGTECGVEKWCGRCSRSLVELTPIAAVHGR 409
Db 534 ---DICKALWCHR---IGRKCTKFMPAEAGTLCQDMWCRGGQCVKYGDEGP-KPTHGH 586
Qy 410 WSWGPRSPCSRSCGGVVTTRRRCNNPRPAFGGRACVGAQLQAEMCNTOACEKTOLEFM 469
Db 587 WSDWSPWSPCSRTCGGISHRDRLCTNPRPSHGGRFCQGSTRTLKLCSQRCPDLSVDFR 646
Qy 470 SQCARTDGGLRSPSGASFYHWGAAPVHQS--GDALCRHMCRAIGESFIMKRGDSFLD 527
Db 647 AAQCAEYNSKFR-----GMLYKW---KPYTQLEDQDLCKLYCIAEGFFFFSLSNKVKD 698
Qy 528 GTRCMPSPGREDGTLISLVSGSCRTFCGDRMDSQQVWDRCCVCGGDNSTCSPRKGSFTA 587
Db 699 GTPC-----SEDSRNVIDGNCERVGCDNVLGSDATEDSCGVCKGNNSDCVTHRGLYSK 752
Qy 588 GRA-REYVTFLTVTNLTYSVIANHRPLFTHLAVRIG-GRYVVGKMSISFNITYPYSLLE 645
Db 753 HHSTNQYHMYTIPSGARSHIYETNISTYSIVSNLSKRYVYLNHWSVD---WP--- 804
Qy 646 DGRVEYRVALTEDRLPRLEIRIWIQLQDADIQVYRRYGEYCNLTRPD-----ITFT 699
Db 805 -GRYKFGSGATFN-----YKRSYKEPENTLSPGPTNETLVEL 840
Qy 700 YFQ-----PKPQAVWAAVRGPCSVSCG 723
Db 841 LFQGRNPGVAWEFSLPRSGAKTTPAAQPSYSWAIVRSECSVSCG 884

RESULT 13
Q7YS95
ID Q7YS95 PRELIMINARY; PRT; 839 AA.
AC Q7YS95;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aggreacanase-1.
GN ADAMTS-4.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OX NCBI_TaxID=9913;

RP	SEQUENCE FROM N.A.	OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
RA	Arai M., Anderson D., Annis B., Collins-Racie L., Corcoran C.,	CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RA	DiBlasio-Smith E., Morris E., Dorner A., LaVallie E.,	OX	NCBI_TaxID=10090;
RT	"Cloning and characterization of bovine aggrecanase-1";	RN	[1]
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.	RP	SEQUENCE FROM N.A.
DR	EMBL; AF515915; AAP47196.1; -.	RA	Strausberg R.,
SQ	SEQUENCE 839 AA; 90279 MW; 1B488A27DF5B96B1 CRC64;	RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
Query Match 9.8%; Score 759.5; DB 6; Length 839;			
Best Local Similarity 31.4%; Pred. No. 2.6e-47;			
Matches 219; Conservative 85; Mismatches 289; Indels 105; Gaps 26;			
QY	61 LQALEPQAVSSYSPGA-----PLKGR-----PPSGFQQRQRRRAAGGI 102	DR	MGI; 1339949; Adamt84.
Db	162 IQLEGAPNSAGPGAGHILRRKSPVSGPGMCMVKAAPPKPS-----SPRAKRFASLS 217	DR	GO; GO:0004222; F:metalloendopeptidase activity; IEA.
QY	103 LHLELIVAGPDVFQAHQEDTERVLTNLNIGAEILLRDPISLCAQFRVHLVKMVLITEPEG 162	DR	GO; GO:0008270; F:zinc ion binding; IEA.
Db	218 RFVETLVADDKMAAFHAGLKYLLTVMAAAKAFKHPISRNPNVSLVTVRLVWLPQEE 277	DR	GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
QY	163 APNITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTQLG 222	DR	GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
Db	278 GQVGFSAQTLRSFCMQRGLNTPDDADPGHFDTAILFTQDGL-CGVSTCDTLGMADVG 336	DR	InterPro; IPR006586; ADAM cysteine.
QY	223 GACSPWTSCLITDGTGDLGVTTIAHEICHSPGLEHAGPSSGC-----GP---SGHVMASD 275	DR	InterPro; IPR001590; Peptidase M12B.
Db	337 TVCDPARSCAIVEDDGLQSAFTAHELGHVFMHLHNS--KQCTGLNGPSTSRHVMAPV 394	DR	InterPro; IPR006025; Pept M_Zn_BS.
QY	276 GAAPRAGLAWSPCSRROLLSLSAGRARCVMDPPRPPQSGSAGHPDPAQFGLYYSANEQCR 335	DR	InterPro; IPR000884; TSP1_1.
Db	395 MAHVDPPEPWPSCSARFITDPLDNGFGHCLLD---KPEAPLHLPTFFPKDYDADRQCC 450	DR	InterPro; IPR008085; TSP1_1.
QY	336 VAFGPXAVACTFAREHLDMAQSLCHTDPLDQSSCSRLLVLLDTEGVEKWCCKGRCR 395	DR	PFam; PF01421; Reprolysin; 1.
Db	451 LTFGPDPSRHCP---QLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCPAQCACMGGRCL 507	DR	PRINTS; PR01705; TSP1REPEAT.
QY	396 SLVEL---TPIAAVHGRWSWSPRSCSRSCGGGVVTRRQCNPNRPAFGGRACVGDAL 451	DR	SMART; SM00209; TSP1; 1.
Db	508 HVDQLQAFNPQAGGFWGSGW---DCRSRSCGGGVQFSSRDCTRPFRNGKCYCEGRR 564	DR	PROSITE; PS00215; ADAM_MEPPO; 1.
QY	452 QAEMCNTQACEK-TQLEFMSQOCA---RTDQQLRSPFGGSAFYHMGAAVPHSQGDA-- 504	DR	PROSITE; PS00092; TSP1; 1.
Db	565 RFRSCNTQDCTGSAITFREQCAAYNHRD--LFPKNFPGMD---W---VRYTYGVAPR 616	QY	223 GACSPWTSCLITDGTGDLGVTTIAHEICHSPGLEHAGP-----GSGCGSPGHVMSADG 276
QY	505 -LCRHWG--RAIGSFIMKRGDSFLDGTGRCMPSPREDGTLSCVSGSCRTFGCDGRMDS 561	Db	333 TVCDPARSCAIVEDDGLQSAFTAHELGHVFMHLHNSKPCNTNLNGQ--GSSRHVMAPVM 391
Db	617 DQCKLTQTRALGYVYVL--DPRVADGTPCSFD-----SSSVCVQGRCIHAGCDRVIGS 668	QY	277 AAPRAGLAWSPCSRROLLSLSAGRARCVMDPPRPPQSGSAGHPDPAQFGLYYSANEQCRV 336
QY	562 QVWDRQVCGGDNSTCSPRKGSFTAGRAREYVTVLTV-----TPNLTSTVI 608	Db	392 AHVDPEEPWPSCSARFITDPLDNGYGHCLLD---KPEAPLHLPTFFPKDYDADRQCC 447
Db	669 KKKFKDKMVCGGDSSCSKQSGSPKKFR-YGYNNVVTIPAGATHILVRQQGSFVRSIYL 727	QY	337 AFGPKAVACTFAREHLDMAQSLCHTDPLDQSSCSRLLVLLDTEGVEKWCCKGRCRS 396
QY	609 ANHRPLFTHLAVRIGRYVVGAKMSISPNTYPSLLEDGRVEYRVALTEDRLPRLEEIRI 668	Db	448 TFGPDSSHCP---QLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCGSSQACMGGRCLH 504
Db	728 ALKLP-----DGIALNGEYTLTPSP--DVVLFGAVSLRYS-----GATAASETLSG 773	QY	397 LVEL---TPIAAVHGRWSWSPRSCSRSCGGGVVTRRQCNPNRPAFGGRACVGDALQ 452
QY	669 WGPLQEDADIQVYRRYGEENLTPRDPITFTYFQPKPR 706	Db	505 VDQLKDFENVQA---GGMGPMGWDGCSRTCGGVQFSSRDCTRPFRNGKCYCEGRRTR 561
Db	774 HGFLABELTQLVL-----VAGNPQNARLRYSEFFVPRPR 806	QY	453 AEMCNTQAC-EKTQLEFMSQOCA---RTDQQLRSPFGGSAFY-HMGAAVPHSQGQDALC 506
RESULT 14			
Q8K384	PRELIMINARY; PRT; 833 AA.	Db	562 FRSCNTENCPHGSALTFRREQCAAYNHRD--LFKSFFGPMDWVPRYTGAQPRDQCKLIC 619
AC	O8K384	QY	507 RHMCRAIGESFIMKRGDSFLDGTGRCMPSPREDGTLSCVSGSCRTFGCDGRMDSQQVMD 566
DT	01-OCT-2002 (TEMBLrel. 22, Created)	Db	620 Q--ARALGYVYVL--PRVADGTPCSFD-----TSSVCVQGRCIHAGCDRIGSKKFD 669
DT	01-OCT-2002 (TEMBLrel. 22, Last sequence update)	QY	567 RCQVCGGDNSTCSPRKGSFTAGRAREYVTVLTVTNLTSTVIYANH-----RPLFTHLAVR 621
DE	Similar to a disintegrin-like and metalloprotease (reprolysin type)	Db	670 KMWVCGGGRCSKQSGSFKKFR-YGYSDVVTIPAGATHILVRQQGSLKSIY---LALK 726
DE	with thrombospondin type 1 motif, 4.		
GN	ADAMTS4.		
OS	Mus musculus (Mouse).		

QY 622 IG-GRYVAGKMSISNTTTPSYLLEDGRVEYRVVALTEDRLPRLEIRIWGPLEQDADIQV 680
DB 727 LSDGVALNGEYTLMPST--DVVLPGAVSLRYS---GATAASETLSGHGFLAQPLTLQV 781
QY 681 YRRYGEYGNLTRPDITFTFYQPKP 705
DB 782 L-----VAGNPQVRLRYSFFVPRP 801

RESULT 15
QB8NJ2
ID QB8NJ2 PRELIMINARY; PRT; 845 AA.
AC QB8NJ2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease.
GN ADAMTS4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA "the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL; AX083534; BAC38944.1; -;
DR MGD; MGI:1339949; Adamts4.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsd_1; 1.
DR PRINTS; PR01705; TSP1REFEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1_1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 845 AA; 91239 MW; B496C3190D1A9225 CRC64;

Query Match 9.7%; Score 748.5; DB 11; Length 845;
Best Local Similarity 30.5%; Pred. No. 1.7e-46;
Matches 209; Conservative 92; Mismatches 303; Indels 81; Gaps 23;

QY 61 LQALEPOAVSYSGAPLKR-----PPSPQFQQRQRAAGCI 102
DB 170 LQLEGGALNAGGPGAGHLRRKSPASSQGPCTVKAPSGSPSP----ISRRTKRFASLS 225
QY 103 LHLELLVAGPDVQAQEDTERVLTNINCAELLRDPISLCAQFRVHLVKMVLITPEG 162
DB 226 RFVETLVADDKMAAFHGTGLKRYLLTVMAAAAKAFKPSIRNPVNLVTRILVILSGQE 285
QY 163 APNITANLTSLLSCVGSQINPEDDTPGHADLVLTITRDELPLDGNRQVRGVTLG 222
DB 286 GPQVGPSAAQTLRFCTWRGLNTPNDSPPDHFDTAILFTRODL-CGVSTCDTLGMADV 344
QY 223 GACSPWTSCLITDTGFDLGVITIAHIGHSFGLHDGAP-----GSGCGPSPGHVNASDG 276
DB 345 TVCDPASCAIVEDDGLQSAFTAHELGHVFNWMLHDSNKPCTNLNGQG-GSSRHVWAPVM 403
QY 277 AAPRAGLAWSCSRQLLSLSAGRARCVWDPPRPPGSGAHPDPAQPLGLIYSANEQCRV 336

DB 404 AHVDPEEPWSPCSARFITDPLNGYGHCLLD-----KPEAPLHLDPATPGDKYDADRQCOL 459
QY 337 AFGPKAVACTFAREHLDMCOALSCHTDQLDQSSCSRLVLLDGTGCGVEKWCCKGRCS 396
DB 460 TFGPDSSHCP---QLPPCAALWCSGHLNHAMCQTKHSPWADGTPCGSSACWGRCLH 516
QY 397 LVLEL-----TPIAAVHGRWSSWSPRSCSCGGVVTTRQCNPRPAFGGRACVAGDLQ 452
DB 517 VDQLKDFNVPOA---GGWGPWGPWGDSCRTCGGVQFSSRDCTRFVPRNGGKYCEGRRT 573
QY 453 AEMCNTQAC-EKTQLFMSQOCA-----RTDQPLRSSPGGASFY-HWGAAYPHSQGDALC 506
DB 574 FRSCNTENCPHGSALTFRBEEQCAAYNHRTD--LFKSPGPMWVPRYTGVAAPRDCKLTC 631
QY 507 RHMCRAGISFIMKRGDSFLDGTCTCMPSGPREDTLSLCVSGSCRTFCGDCGRMDSQQVMD 566
DB 632 Q--ARALGYVYLE--PRVADGTPCSPD-----TSSVCVQGRCHAGCDRIIGSKKFD 681
QY 567 RCQVCGGDNSTCSPRKGSFTAGRAREYVTELTVPNLTSVVIANH-----RPLFTHLAVR 621
DB 682 KMWVCGDGRCSKQSGSPKKPR-YGYSDVVTIPAGATHILVRQQGGSGLSIY--LALK 738
QY 622 IG-GRYVAGKMSISNTTTPSYLLEDGRVEYRVVALTEDRLPRLEIRIWGPLEQDADIQV 680
DB 739 LSDGVALNGEYTLMPST--DVVLPGAVSLRYS---GATAASETLSGHGFLAQPLTLQV 793
QY 681 YRRYGEYGNLTRPDITFTFYQPKP 705
DB 794 L-----VAGNPQVRLRYSFFVPRP 813

Search completed: March 13, 2004, 07:43:13
Job time : 147.125 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 207.238 Seconds
(without alignments)
1930.574 Million cell updates/sec

Title: US-09-836-712-2
Perfect score: 7751
Sequence: 1 PGRPTRKAPSHSAPLLGLA.....LQSWPEMQDPSQWKGKGT 1416

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7751	100.0	1416	5	AB04153 Human ADA
2	7604	98.1	1427	5	Aae24450 Human Von
3	7596	98.0	1427	7	Add94038 Human agg
4	7274	93.8	1445	5	Abb98125 Human PMM
5	7190	92.8	1353	5	Aae24449 Human Von
6	7190	92.8	1353	6	Aao16617 Human von
7	7081	91.4	1297	6	Aao16618 Human von
8	6861.5	88.5	1378	6	Aao16619 Human von
9	6752.5	87.1	1323	6	Aao16620 Human von
10	4434.5	57.2	933	5	Aau79217 Human ADA
11	3533.5	45.8	1120	4	Ag63829 Amino aci
12	3533.5	45.8	1120	6	Abg74113 Human mat
13	2340	30.2	984	6	Abg74113 Human NOV
14	1799	23.2	364	5	Aau97642 Human Hsa
15	1799	23.2	364	5	Abp43989 Procollag
16	1799	23.2	365	7	Add94035 Human Hsa
17	1459.5	18.8	738	5	Aau97643 Human Hsa
18	1459.5	18.8	738	7	Add94036 Human Hsa
19	1261	16.3	242	5	Aau97641 Human agg
20	1261	16.3	242	7	Add94031 Human agg
21	1192.5	15.4	1221	6	Aae38366 Human agg
22	1190.5	15.4	1049	7	Add22497 ADAMTS fa
23	1190.5	15.4	1221	5	Abb81460 Human agg
24	1190.5	15.4	1221	7	Add22493 ADAMTS fa
25	1190.5	15.4	1221	7	Add22495 ADAMTS fa

26	1190	15.4	1505	5	AAU72897	Human met
27	1185	15.3	223	4	ABU53113	Human tes
28	1185	15.3	223	5	ABG66717	Human nov
29	1185	15.3	934	5	AAU72893	Human met
30	1176.5	15.2	1629	5	ABG30703	Human agg
31	1176.5	15.2	1629	5	AAO14448	Human ADA
32	1176.5	15.2	1935	5	AAU72896	Human met
33	1173	15.1	1363	7	ADD44592	Polypepti
34	1172.5	15.1	1916	5	Aae19173	Human pro
35	1165.5	15.0	1602	5	ABG30702	Human agg
36	1157	14.9	1882	4	AAU72896	Human ADA
37	1155	14.9	1934	4	AAU72301	Human ADA
38	1148.5	14.8	1224	5	AAU72891	Human met
39	1148.5	14.8	1224	5	AAO15254	Human MDT
40	1148.5	14.8	1224	5	AAU79497	Human ADA
41	1148.5	14.8	1224	6	ABG72431	Novel hum
42	1148.5	14.8	1224	6	ABR83432	Human ADA
43	1148.5	14.8	1224	7	ADA50762	Novel hum
44	1147.5	14.8	1170	5	AAU79500	Human ADA
45	1147.5	14.8	1224	5	AAU80153	Human ADA

ALIGNMENTS

RESULT 1
ABB04153
ID ABB04153 standard; protein; 1416 AA.
XX ABB04153;
XX AC ABB04153;
XX XX

26-MAR-2002 (first entry)
Human ADAMTS-M polypeptide.
XX XX
XX XX

Osteoarthritis; rheumatoid arthritis; inflammatory bowel disease;
Crohn's disease; asthma; Alzheimer's disease; organ transplant rejection;
cachexia; allergy; cancer; leukaemia; lymphoma; osteoporosis;
atherosclerosis; congestive heart failure; myocardial infarction; stroke;
neurodegenerative disease; autoimmune disorder; Huntington's;
Parkinson's; migraine; pain; depression; multiple sclerosis; burn;
infertility; diabetic shock; gene therapy; ADAMTS-M;
A Disintegrin And Metalloprotease; thrombospondin domain.
XX OS Homo sapiens.
XX XX

Key Location/Qualifiers
Peptide 1..97
FT /label= Prodomain
FT /note= "The mature form of the ADAMTS-M protein is
processed by furin cleavage of the prodomain"
FT 94..97
FT Cleavage-site
FT /label= Furin_cleavage_site
FT Protein 98..1416
FT /label= Mature ADAMTS-M protein
FT /note= "The mature form of the ADAMTS-M protein is
processed by furin cleavage of the prodomain"
FT 98..311
FT Domain
FT /label= Metalloprotease_domain
FT 247..272
FT Domain
FT /label= Zinc-binding_motif
FT 324..394
FT Domain
FT /label= Disintegrin_domain
FT 410..473
FT Domain
FT /label= Thrombospondin_submotif
FT 419..424
FT Domain
FT /label= Heparin-binding_domain
FT 1099..1156
FT Domain
FT /label= Thrombospondin_submotif

XX EP1152055-A1.
PN
XX
PD 07-NOV-2001.

XX 24-APR-2001; 2001EP-00303706.
XX 27-APR-2000; 2000US-0200040P.
XX (PIZ.) PRIZER PROD INC.
XX Buckbinder L, Mitchell PG, Wachtmann TS, Walsh RT;
XX WPI; 2002-084275/12.
XX N-PSDB; ABA02549.
XX New polynucleotide, useful in gene therapy, particularly for treating or
XX preventing e.g. arthritis, Crohn's disease, Alzheimer's disease and organ
XX transplant toxicity and rejection, comprises ADAMTS polynucleotide and
XX encoded polypeptide.
XX Claim 4; Fig 2; 31pp; English.
XX
XX The present sequence represents a ADAMTS protein, designated ADAMTS-M,
XX that exhibits the characteristics of the ADAM (A Disintegrin And
XX Metalloprotease) family of metalloproteases, and contains a
XX thrombospondin domain (TS). The protein is encoded by the cDNA given in
XX ABA02549. The specification describes a newly isolated polynucleotide,
XX comprising a nucleotide sequence encoding an ADAMTS-M polypeptide as
XX given in the specification, or a metalloproteinase, disintegrin domain,
XX prodomain or its thrombospondin submotif. The polynucleotide, polypeptide
XX and agent are useful for manufacturing a medicament for treating a
XX subject in need of altering activity or expression of ADAMTS-M. The
XX polynucleotide, ADAMTS-M polypeptide and agent are useful for
XX manufacturing a medicament for treating arthritis (osteoarthritis and
XX rheumatoid arthritis), inflammatory bowel disease, Crohn's disease,
XX asthma, Alzheimer's disease, organ transplant toxicity and rejection,
XX cachexia, allergy, cancer (e.g. solid tumour cancer including colon,
XX breast, lung, prostate, brain or haematopoietic malignancies including
XX leukaemia and lymphoma), osteoporosis, atherosclerosis, aortic aneurysm,
XX congestive heart failure, myocardial infarction, stroke, head trauma,
XX spinal cord injury, neurodegenerative disease, autoimmune disorders,
XX Huntington's disease, Parkinson's disease, migraine, pain, depression,
XX multiple sclerosis, abnormal wound healing, burns, infertility or
XX diabetic shock. The polynucleotide and polypeptide are also useful for
XX diagnosing the diseases above. The polynucleotide is particularly useful
XX in gene therapy for treating the diseases cited above
XX
XX Sequence 1416 AA;

QY	361	HTDPLDQSSRLIVPLLDGTEGVEKWCCKGRCSRLVELTPIAAVHGRWSSNGRSPCS	420
Db	361	HTDPLDQSSRLIVPLLDGTEGVEKWCCKGRCSRLVELTPIAAVHGRWSSNGRSPCS	420
QY	421	RSCGGVTVRRQCNPRPAFGACVADLOEMCMTQACEKTQLEFMSQCCARTDQOP	480
Db	421	RSCGGVTVRRQCNPRPAFGACVADLOEMCMTQACEKTQLEFMSQCCARTDQOP	480
QY	481	LRSPGGASFYHWAAGVPHSQDQALCRHMCRAIGESFIMKRGDSFLDGTGTRCMPSPREDG	540
Db	481	LRSPGGASFYHWAAGVPHSQDQALCRHMCRAIGESFIMKRGDSFLDGTGTRCMPSPREDG	540
QY	541	TLSLCVSGSCRTFGCDGRMDSQVWDRQVCGGDNSTCSPRKGSFTAGRAREYTFLLTVT	600
Db	541	TLSLCVSGSCRTFGCDGRMDSQVWDRQVCGGDNSTCSPRKGSFTAGRAREYTFLLTVT	600
QY	601	PNLTSVYIANHREPLFTHLAVRIGGRVYVAGKMSISPNNTYPSLLEDGVEYRVVALTEDRL	660
Db	601	PNLTSVYIANHREPLFTHLAVRIGGRVYVAGKMSISPNNTYPSLLEDGVEYRVVALTEDRL	660
QY	661	PRLEIRIWGPLOEDADIQVYRYGEEYGNLTPDITFTYFQPKPRQAWVAAVVGPCSV	720
Db	661	PRLEIRIWGPLOEDADIQVYRYGEEYGNLTPDITFTYFQPKPRQAWVAAVVGPCSV	720
QY	721	SCGAGLRWVNYSCLDQARKELVETVQCQSQQPPANPEACVLEPCPPYVAVGDFGPCSAS	780
Db	721	SCGAGLRWVNYSCLDQARKELVETVQCQSQQPPANPEACVLEPCPPYVAVGDFGPCSAS	780
QY	781	CGGRLRERVRVCEAAGSLLKTLPPARCAGAAQPAVALETNCNPPCARVEVEPSSCT	840
Db	781	CGGRLRERVRVCEAAGSLLKTLPPARCAGAAQPAVALETNCNPPCARVEVEPSSCT	840
QY	841	SAGGAGLALENETCVFAGDLEAPVTEGPGSVDEKLPAPPCVGMSPGPGHGLDTSAG	900
Db	841	SAGGAGLALENETCVFAGDLEAPVTEGPGSVDEKLPAPPCVGMSPGPGHGLDTSAG	900
QY	901	EKAPSPWGSIRTGAAAHVWTPAAGSCSVSCGRGLMELRFLCMDLSALRVPVQEEELCGLAS	960
Db	901	EKAPSPWGSIRTGAAAHVWTPAAGSCSVSCGRGLMELRFLCMDLSALRVPVQEEELCGLAS	960
QY	961	KFGSRREVCOAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTQCG	1020
Db	961	KFGSRREVCOAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTQCG	1020
QY	1021	LPRPQEAACSLPCPPRKKWMSLGPSCSASCGLGTARRSVACVQDDQDVEVEAACA	1080
Db	1021	LPRPQEAACSLPCPPRKKWMSLGPSCSASCGLGTARRSVACVQDDQDVEVEAACA	1080
QY	1081	LVRPEASVCLTADCTYRWHVGTWMECSVSCDGIQRRRDTCLGPAQAPVADFCQHL	1140
Db	1081	LVRPEASVCLTADCTYRWHVGTWMECSVSCDGIQRRRDTCLGPAQAPVADFCQHL	1140
QY	1141	KPVTVRGCVAGPCVGGTSPSLVPHHEAAAFGRITATPAGACGRQHLEPTGTIDMRGPQA	1200
Db	1141	KPVTVRGCVAGPCVGGTSPSLVPHHEAAAFGRITATPAGACGRQHLEPTGTIDMRGPQA	1200
QY	1201	DCAVAIGRPLGEVTVLRLVLESSLNCAGDMLLWGRLTWRKMKRLLDNTFSSKNTLVV	1260
Db	1201	DCAVAIGRPLGEVTVLRLVLESSLNCAGDMLLWGRLTWRKMKRLLDNTFSSKNTLVV	1260
QY	1261	RORCGPGGVLLRYGSQLAPETFYRECWMQLEGPWGEIYVSPISPATNSAGGCRFLINV	1320
Db	1261	RORCGPGGVLLRYGSQLAPETFYRECWMQLEGPWGEIYVSPISPATNSAGGCRFLINV	1320
QY	1321	APHARIAIHALATNMAGTGANASYILIRDTLSLRTAFHGQOVLVWESSSSQAEEMBFS	1380
Db	1321	APHARIAIHALATNMAGTGANASYILIRDTLSLRTAFHGQOVLVWESSSSQAEEMBFS	1380
QY	1381	EGFLKAQASLRGQWTLQSWVPEMQDPQSWKKEGT 1416	
Db	1381	EGFLKAQASLRGQWTLQSWVPEMQDPQSWKKEGT 1416	

RESULT 2	
AAE24450	
ID	AAE24450 standard; protein; 1427 AA.
XX	AC
XX	AAE24450;
XX	DT
XX	04-OCT-2002 (first entry)
XX	DE
XX	Human Von Willebrand factor-cleaving protease (vWF-cp).
XX	Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;
KW	transgenic animal; immunisation; thromboembolic disease; preeclampsia;
KW	thrombotic thrombocytic purpura; TTP; Henoch-Schonlein purpura;
KW	thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;
KW	transgenic; anticoagulant; chromosome 9.
XX	OS
XX	Homo sapiens.
XX	PH
XX	Key
XX	Location/Qualifiers
XX	1..51
XX	/label= Signal_peptide
XX	52..1427
XX	/note= "Mature human vWF-cp protein"
XX	69..75
XX	/note= "Purin cleavage site"
XX	224..228
XX	/note= "Catalytical side"
XX	249
XX	/note= "Met turn"
XX	301..377
XX	/note= "Distintegrin like motif"
XX	387..439
XX	/note= "Thromspondin type I motif"
XX	441..553
XX	/note= "Cysteine rich region"
XX	554..687
XX	/note= "Spacer"
XX	688..743
XX	/note= "Thromspondin type I motif"
XX	744..805
XX	/note= "Thromspondin type I motif"
XX	897..952
XX	/note= "Thromspondin type I motif"
XX	953..1013
XX	/note= "Thromspondin type I motif"
XX	1016..1073
XX	/note= "Thromspondin type I motif"
XX	1075..1131
XX	/note= "Thromspondin type I motif"
XX	WO200242441-A2.
XX	30-MAY-2002.
XX	20-NOV-2001; 2001WO-EP013391.
XX	22-NOV-2000; 2000US-00721254.
XX	12-APR-2001; 2001US-00833328.
XX	(BAXT) BAXTER AG.
XX	Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
PI	Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
PI	Zimmermann K, Voelkel D;
XX	WPI; 2002-479950/51.
DR	N-PSDB; AAD39332.
XX	Novel isolated or substantially purified Von Willebrand factor-cleaving
PT	protease, useful for producing preparation for therapy of thrombosis and
PT	thromboembolic disease such as thrombotic thrombocytic purpura.
XX	
PS	Claim 1; Fig 5; 93pp; English.
XX	The invention relates to an isolated or substantially pure Von Willebrand
CC	factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
CC	purifying vWF which involves providing vWF-cp as a ligand, contacting a
CC	solution comprising vWF with the polypeptide ligand under conditions
CC	where vWF is bound to the ligand and recovering from the ligand purified
CC	vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
CC	which involves immunising an animal with vWF-cp and isolating the anti-
CC	vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
CC	producing a preparation of prophylaxis and therapy of thrombosis and
CC	thromboembolic disease such as thrombotic thrombocytic purpura (TTP),
CC	Henoch-Schonlein purpura, preeclampsia, neonatal thrombocytopenia or
CC	haemolytic-uraemic syndrome. vWF-cp can also be used for processing
CC	plasmatic or recombinantly produced vWF. The invention is useful for
CC	construction of recombination systems and generating transgenic animals which
CC	express the polypeptide in vivo. The present sequence is human vWF-cp
CC	protein. vWF-cp gene is located on chromosome 9
XX	Sequence 1427 AA;
SQ	
	Query Match 98.1%; Score 7604; DB 5; Length 1427;
	Best Local Similarity 97.6%; Pred. NO. 0;
	Matches 1393; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
QY	24 MEQRHPRARCPPLCVAGILACGFLLCWGPSHFQSCLOALEPQAVSSYLSFGAPLKGRP 83
DB	1 MEQRHPRARCPPLCVAGILACGFLLCWGPSHFQSCLOALEPQAVSSYLSFGAPLKGRP 60
QY	84 PSPGFQORQORRAAGGILHLELLVAVGDPVQAHQEDTERVYLTNINIGAEILRDPSL 143
DB	61 PSPGFQORQORRAAGGILHLELLVAVGDPVQAHQEDTERVYLTNINIGAEILRDPSL 120
QY	144 GQAFRVHLVKWVILTEPEGAPNITANLTSSLLSCVGSQNTINPEDDTPGHADLVLYITR 203
DB	121 GQAFRVHLVKWVILTEPEGAPNITANLTSSLLSCVGSQNTINPEDDTPGHADLVLYITR 180
QY	204 FDLPLDGNRQVRGVTQLGACSPTWSCLTITDGTDLGVYITAEIHGHSFGLHGDGAPGS 263
DB	181 FDLPLDGNRQVRGVTQLGACSPTWSCLTITDGTDLGVYITAEIHGHSFGLHGDGAPGS 240
QY	264 GCGPSGHVNASDGAAPRAGLAWSPCSRRLILLSAGRCVWDPDPQPSAGHPDDAQ 323
DB	241 GCGPSGHVNASDGAAPRAGLAWSPCSRRLILLSAGRCVWDPDPQPSAGHPDDAQ 300
QY	324 PGLYTSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTETC 383
DB	301 PGLYTSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTETC 360
QY	384 GVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGG 443
DB	361 GVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGG 420
QY	444 RACVGADLOAEMCNTQACEKTQLEFMSQCCARTDQQLRSPGGASFYHWGAAPHSQGD 503
DB	421 RACVGADLOAEMCNTQACEKTQLEFMSQCCARTDQQLRSPGGASFYHWGAAPHSQGD 480
QY	504 ALCRHMCRAIGESFIMKRGDSFLDGTFCMPSGREDGTLSLCVSGSCRTFCGCDGRMDSQ 563
DB	481 ALCRHMCRAIGESFIMKRGDSFLDGTFCMPSGREDGTLSLCVSGSCRTFCGCDGRMDSQ 540
QY	564 VMDRCVCGDNDSTCSPRKGSEPTAGRAREVYVTLTVPNTLSVYIANHRPLFTHLAVRIG 623
DB	541 VMDRCVCGDNDSTCSPRKGSEPTAGRAREVYVTLTVPNTLSVYIANHRPLFTHLAVRIG 600
QY	624 GRYVWAGKMSISPNTTYPSSLEDDGRVEYRVALTEDRLPRLEIRIWGPLEQEDADIQVYRR 683
DB	601 GRYVWAGKMSISPNTTYPSSLEDDGRVEYRVALTEDRLPRLEIRIWGPLEQEDADIQVYRR 660
QY	684 YGEYGNLTRPDITFTFYQPKPQAWVAAVRGPCSVSCGAGLRWVNTSCLDQARKELVE 743
DB	661 YGEYGNLTRPDITFTFYQPKPQAWVAAVRGPCSVSCGAGLRWVNTSCLDQARKELVE 720

PR	16-OCT-2001; 2001IUS-00978797.
XX	(AMHP) AMERICAN HOME PROD CORP.
PA	
XX	
XX	Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
XX	
DR	WPI: 2003-801251/75.
DR	N-PSDB; AD94037.
XX	
PPT	New isolated DNA molecule encoding an aggrecanase polypeptide for producing a purified human aggrecanase protein which can be used to develop inhibitors of aggrecanase.
PT	
XX	
XX	Claim 14; SEQ ID NO 8; 24pp; English.
XX	
CC	This invention relates to a novel human aggrecanase protein and the DNA sequence which encodes it. Aggrecan is a major extracellular component of articular cartilage. It is a proteoglycan responsible for providing cartilage with its mechanical properties of compressability and elasticity. A proteolytic activity (aggrecanase) is responsible for the cleavage of aggrecan thereby having a role in cartilage degradation associated with osteoarthritis and inflammatory joint disease. Compounds which inhibit the activity of the protein of the invention may have antiarthritic, osteopathic or antiinflammatory activity. The invention may be used to produce a purified human aggrecanase protein. The protein (or fragment) may be used to develop inhibitors of aggrecanase, using three dimensional structural analysis or computer aided drug design. A peptide which binds to aggrecanase is used to inhibit the proteolytic degradation of aggrecan. The invention may be useful for the development of therapeutics for the treatment of aggrecanase-associated disorders, such as, osteoarthritis and other inflammatory conditions. The present sequence is that of the human aggrecanase protein (full length sequence) of the invention.
CC	
XX	
SQ	Sequence 1427 AA;
Query Match	98.0%; Score 7596; DB 7; Length 1427;
Best Local Similarity	97.5%; Pred. No. 0;
Matches 1392; Conservative	0; Mismatches 1; Indels 34; Gaps 1;
QY	24 MHQRHPRARCPPLCVAGILACGFLLCWGSPHFQQSCLOALEPOAVSSVLSPCAPLKGRRP 83
Db	1 MQQRHWARCPPLCVAGILACGFLLCWGSPHFQQSCLOALEPOAVSSVLSPCAPLKGRRP 60
QY	84 PSPGFQQRQRARRAGGIHLLELLVAVGPDVFQAHCEDTERLYLTNLNIGAEILLRDPSL 143
Db	61 PSPGFQQRQRARRAGGIHLLELLVAVGPDVFQAHCEDTERLYLTNLNIGAEILLRDPSL 120
QY	144 GQAQFRVHLVKMVLITPEGAPNITANLTSSLISVCWSQTINPEDDTDPGHADLVLYITR 203
Db	121 GQAQFRVHLVKMVLITPEGAPNITANLTSSLISVCWSQTINPEDDTDPGHADLVLYITR 180
QY	204 FDLLELDGNRGVRGVTLGGACSPTWSCSLITEGTDFLGVTIAHEIGHSGFLEHDGAFGS 263
Db	181 FDLLELDGNRGVRGVTLGGACSPTWSCSLITEGTDFLGVTIAHEIGHSGFLEHDGAFGS 240
QY	264 GCGPSGHVNASDGAAPRAGLAWSPCSRROLLSLSAGRARCVDMPRPQPGSGAGHPDDAQ 323
Db	241 GCGPSGHVNASDGAAPRAGLAWSPCSRROLLSLSAGRARCVDMPRPQPGSGAGHPDDAQ 300
QY	324 PGLYYSANECQRFVAFGPKAVACTFAREHLDMCQALSCHTDPDQSSCSRLIVLPDLDTGTC 383
Db	301 PGLYYSANECQRFVAFGPKAVACTFAREHLDMCQALSCHTDPDQSSCSRLIVLPDLDTGTC 360
QY	384 GVEKWCCKGRCLSLVELTPTAHVHGWSWGPRSPCSRSCGGVTVTRRCNNPRPFAFGG 443
Db	361 GVEKWCCKGRCLSLVELTPTAHVHGWSWGPRSPCSRSCGGVTVTRRCNNPRPFAFGG 420
QY	444 RACVGADLOAEMCNVTOACEKTQLFEFMSQCARTDGPRLSRPPGASFYHHGAAVPHSQGD 503
Db	421 RACVGADLOAEMCNVTOACEKTQLFEFMSQCARTDGPRLSRPPGASFYHHGAAVPHSQGD 480
QY	504 ALCRHMCAIGESFIWKRGDSFLDGTTCMPSPGPREDTGLSLCVSGSCRITFGDGRMDSSQ 563

Db 481 ALCRHMCRAIGESFIMKRGDSFLDGTFCMSPREDGTLSCVSGSCHTFFGDCGRMDSQQ 540
QY 564 VVDRCOVCGDNDSCSPKRGSTAGRAREVYVTLVTPNLTSVVIANRPLFTHLAVRIG 623
Db 541 VVDRCOVCGDNDSCSPKRGSTAGRAREVYVTLVTPNLTSVVIANRPLFTHLAVRIG 600
QY 624 GRVYVAGKWSISNTTYPSPLELGRVYRVVALTEDRLPRLEIRIINGWLPQEDADIQVYR 683
Db 601 GRVYVAGKWSISNTTYPSPLELGRVYRVVALTEDRLPRLEIRIINGWLPQEDADIQVYR 660
QY 684 YGEEYGNLTRPDITFTYFQPKPQAAVWAAVGRPCSVSCGAGLAWNVYSCLDQARKELVE 743
Db 661 YGEEYGNLTRPDITFTYFQPKPQAAVWAAVGRPCSVSCGAGLAWNVYSCLDQARKELVE 720
QY 744 TVQCGSQSQPPAWPEACVLEPCPPYVAVDFGPCSASCGGGLRPRVRCVBAQGSLLKTL 803
Db 721 TVQCGSQSQPPAWPEACVLEPCPPYVAVDFGPCSASCGGGLRPRVRCVBAQGSLLKTL 780
QY 804 PPARCRAGAQQPAVALTECNQPCPARWEVSEPSCTSSAGAGLALENETCVPGADGLEA 863
Db 781 PPARCRAGAQQPAVALTECNQPCPARWEVSEPSCTSSAGAGLALENETCVPGADGLEA 840
QY 864 PVTGPGSVDEKLPAPPCVGMSCPPGWHLDATSAKEKAPSPWGSIRTGAAHVWTPA 923
Db 841 PVTGPGSVDEKLPAPPCVGMSCPPGWHLDATSAKEKAPSPWGSIRTGAAHVWTPA 900
QY 924 AGSCSVSCGRGLMELFLCMLSALRVVQVBEELCLASKPGSRREVCOAVPCPARWQYKLA 983
Db 901 AGSCSVSCGRGLMELFLCMLSALRVVQVBEELCLASKPGSRREVCOAVPCPARWQYKLA 960
QY 984 ACSVSCGRGVVRRILYCARAHGDDGEEIILDTQCQGLPRPEQOEACSLFPCPPRWKMS 1043
Db 961 ACSVSCGRGVVRRILYCARAHGDDGEEIILDTQCQGLPRPEQOEACSLFPCPPRWKMS 1020
QY 1044 LGPCASCSGLGTARRSVACVOLQGGQVDEVEDEAAALVRPEASVPCLIADCTRWVHVT 1103
Db 1021 LGPCASCSGLGTARRSVACVOLQGGQVDEVEDEAAALVRPEASVPCLIADCTRWVHVT 1080
QY 1104 WMECSVSCDGIQRRDTCGLGPOAQAPVADFCOHLKPVTVRGCGWAGPCVGGTSPSLVP 1163
Db 1081 WMECSVSCDGIQRRDTCGLGPOAQAPVADFCOHLKPVTVRGCGWAGPCVGGTSPSLVP 1140
QY 1164 HEEAAPGRTTATPAG-----ACGRQHLEPT 1189
Db 1141 HEEAAPGRTTATPAGASLEWSQAGLLFSPAPQPRLLPGQNSVQSSACGRQHLEPT 1200
QY 1190 GTIDMEGPGQADCAVAIGRELGEVTVLRVLESSLNCAGDMMLLWGLRTWRKCRKLLDM 1249
Db 1201 GTIDMEGPGQADCAVAIGRELGEVTVLRVLESSLNCAGDMMLLWGLRTWRKCRKLLDM 1260
QY 1250 TFSSKNTLVVRCQGRPGGVLLRYGSQLAPETFYRECDNQLFGPWGEIVSPSLSPATS 1309
Db 1261 TFSSKNTLVVRCQGRPGGVLLRYGSQLAPETFYRECDNQLFGPWGEIVSPSLSPATS 1320
QY 1310 NAGGCLFNVAPHARIAHALATNMAGTEGANASYILIRDTLSLRTAFHQOQVLYWE 1369
Db 1321 NAGGCLFNVAPHARIAHALATNMAGTEGANASYILIRDTLSLRTAFHQOQVLYWE 1380
QY 1370 SSSSQAEFFSEGLKAAQSLRGQYWTQLQSWVPMQDPQSNKKEGT 1416
Db 1381 SSSSQAEFFSEGLKAAQSLRGQYWTQLQSWVPMQDPQSNKKEGT 1427

RESULT 4

ABB98125

ID ABB98125 standard; protein; 1445 AA.

XX AC ABB98125;

XX DT 17-OCT-2002 (first entry)

XX DE Human PMM Incyte ID 7473607CD1.

XX KW Human; PMM; protein modification and maintenance molecule;
KW anticonvulsant; neuroprotective; nootropic; cytoskeletal; antipsoriatic;
KW antineoplastic; dermatological; antidiabetic; antiparkinsonian;
KW antianemic; antiinflammatory; antiulcer; antitubercular; cardiac;
KW hepatotropic; osteopathic; antiemetic; antipruritic; virucide;
KW antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
KW haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
KW cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
KW antitumor; antirheumatic; immunosuppressive; antiallergic; antithyroid;
KW nephrotropic; angiotensin; thymimetic; antidiarrhetic; uropathic;
KW ophthalmological; antiparasitic; tranquilizer; vulnerary; keratolytic;
KW auditory; antiseborrheic; antidepressant; neuroleptic; antinfertility;
KW anthelmintic; protozoacide; Crohn's disease; hypertension; autoimmune;
KW inflammatory; aneemia; cell proliferative; developmental; epithelial;
KW scabies; neurological; Alzheimer's disease; reproductive;
KW ectopic pregnancy; Gene therapy; vaccine; disorder;
KW procollagen I N-proteinase.
XX Homo sapiens.
OS
XX WO200246383-A2.
PN
XX 13-JUN-2002.
XX
XX 05-DEC-2001; 2001WO-US046964.
XX
XX 08-DEC-2000; 2000US-0254399P.
PR
XX 21-DEC-2000; 2000US-0257803P.
PR
XX 05-JAN-2001; 2001US-0260110P.
PR
XX 19-JAN-2001; 2001US-0262851P.
PR
XX 25-JAN-2001; 2001US-0264623P.
PR
XX (INCY-) INCYTE GENOMICS INC.
PA
XX Yue H, Azimzai Y, Kallick DA, Baughn MR, Griffin JA, Swarnakar A;
PI Lal PG, Wallia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;
PI Ramkumar J, Thangavelu K, Iu Y, Warren BA, Lu DAM, Lee EA;
PI Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;
PI Sanjanwala MV;
XX WPI; 2002-519664/55.
DR N-PSDB; ABQ75946.
XX
XX New isolated Protein Modification and Maintenance polypeptides, useful
PT for diagnosis, and treatment of e.g. gastrointestinal disorders.
XX
XX Claim 1 (a); Page 154-157; 200pp; English.
XX
XX The invention relates to an isolated Protein Modification and Maintenance
CC polypeptide. Polypeptides of the invention may be used in the
CC diagnosis, treatment and prevention of disorders associated with
CC decreased expression or activity of PMM. These include gastrointestinal
CC disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
CC hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
CC proliferative disorders, developmental disorders (e.g. Alzheimer's disease)
CC (e.g. scabies), neurological disorders (e.g. ectopic pregnancy), and in gene therapy or a
CC vaccine for such diseases. They may also be used in the assessment of the
CC effects of exogenous compound on the expression of nucleic acid and amino
CC acid sequences of protein modification and maintenance molecules. The
CC current sequence represents a human PMM of the invention, which has been
CC found to have homology with e.g. taurine procollagen I N-proteinase
XX
SQ Sequence 1445 AA;

Query Match 93.8%; Score 7274; DB 5; Length 1445;
Best Local Similarity 96.1%; Pred. No. 0;
Matches 1338; Conservative 0; Mismatches 4; Indels 50; Gaps 2;

QY 74 SPGAPLKGRPPSPGQFQRRQRRAGGILHLELLVAVGPDVFOAQHEDTERVLTNMI 133
DB 55 SPSPFP-PGRPPSPGQFQRRQRRAGGILHLELLVAVGPDVFOAQHEDTERVLTNMI 113

134 GAELLRDPGLGAQFRVHLVQWVILTEPEGAPNITANLTSSLLSVCGNSQTNIPEDDTPG 193
114 GAELLRDPGLGAQFRVHLVQWVILTEPEGAPNITANLTSSLLSVCGNSQTNIPEDDTPG 173
194 HADLVLYITRFDLELPGNRQVRGVTQLGGACSPSTWCLITETGDFDLGVIAHEIGHSF 253
174 HADLVLYITRFDLELPGNRQVRGVTQLGGACSPSTWCLITETGDFDLGVIAHEIGHSF 233
254 GLEHDGAPGSGCPGSHVMSDGAAPRAGLAWSPCSRRLSLLSAGPARCWDPPRPQP 313
234 GLEHDGAPGSGCPGSHVMSDGAAPRAGLAWSPCSRRLSLLSAGPARCWDPPRPQP 293
314 GSAGHPDPAQGLIYSANEOCRVAFGKAVACTFARHLDNCOALSCHTDPLDQSSCSRL 373
294 GSAGHPDPAQGLIYSANEOCRVAFGKAVACTFARHLDNCOALSCHTDPLDQSSCSRL 353
374 LVPLLDGTGCGVKWCKGCRSLVELTPIAAVHGRSSWGPRSPCSRSGGVVTRRQ 433
354 LVPLLDGTGCGVKWCKGCRSLVELTPIAAVHGRSSWGPRSPCSRSGGVVTRRQ 413
434 CNNPRAFGGRACVADLQAEKQCEKQLEBMSQCCARTDQQLRSPGGASFYHW 493
414 CNNPRAFGGRACVADLQAEKQCEKQLEBMSQCCARTDQQLRSPGGASFYHW 473
494 GAAVPHSQDALCRHMCRAIGESFIMKRGDSFLDTRCWPSPGPRDGLSLCVSGSRTF 553
474 GAAVPHSQDALCRHMCRAIGESFIMKRGDSFLDTRCWPSPGPRDGLSLCVSGSRTF 533
554 GCDGRMDSQWDRQCQCGGDNSTCSRKGSFTAGAREYVTFLTNTSVYIANHRP 613
534 GCDGRMDSQWDRQCQCGGDNSTCSRKGSFTAGAREYVTFLTNTSVYIANHRP 593
614 LFTHLAVRIGRYVAVGKMSISPNVTYPSLLEDGRVEYRVVALTEDRLEIRIWGPLO 673
594 LFTHLAVRIGRYVAVGKMSISPNVTYPSLLEDGRVEYRVVALTEDRLEIRIWGPLO 653
674 EDADIQVYRYGEBYGNLTPTDITFTYQPKPQAWVMAAVRGPCSVSCGAGLRVWVNSC 733
654 EDADIQVYRYGEBYGNLTPTDITFTYQPKPQAWVMAAVRGPCSVSCGAGLRVWVNSC 713
734 LDQARKELVTVQCGSQQPAPWAEACVLEPCPPYVAVDGFPCASCGGLRERPRVCV 793
714 LDQARKELVTVQCGSQQPAPWAEACVLEPCPPYVAVDGFPCASCGGLRERPRVCV 773
794 EAQSLKLTLPAPCRAGAQPAVALETCNPPCPARWEVSEPSCTSSAGGAGLAENET 853
774 EAQSLKLTLPAPCRAGAQPAVALETCNPPCPARWEVSEPSCTSSAGGAGLAENET 833
854 CVPAGDGLAPVTEGPGSVDEKLPAPEPCVGVSCPPGWHLDATGAGEKAPSPWGSIRTG 913
834 CVPAGDGLAPVTEGPGSVDEKLPAPEPCVGVSCPPGWHLDATGAGEKAPSPWGSIRTG 893
914 AQAAHVWTPAAGSCSVSCGRGIMELRFLCMLSALRVPVQBELCGLASKPGSRREVCQAVP 973
894 AQAAHVWTPAAGSCSVSCGRGIMELRFLCMLSALRVPVQBELCGLASKPGSRREVCQAVP 953
974 CPARWQYKLAACSVSGRGVVRILYCARAGEDDGEIILLDTQCGLPRPBPQACSLIE 1033
954 CPARWQYKLAACSVSGRGVVRILYCARAGEDDGEIILLDTQCGLPRPBPQACSLIE 1013
1034 PCPPRWKVMISLGPSCASCLGTARRSVACVLDQGDVEVDEAAACALVRPEASVPCLIA 1093
1014 PCPPRWKVMISLGPSCASCLGTARRSVACVLDQGDVEVDEAAACALVRPEASVPCLIA 1073
1094 DCTVRHVGTWMECSVSCGDIQRRDTCGLGPAQAPVADFCQHLPKPVTVRGCGWAGPC 1153
1074 DCTVRHVGTWMECSVSCGDIQRRDTCGLGPAQAPVADFCQHLPKPVTVRGCGWAGPC 1133
1154 VGQGTPLVPHEEAAAPGRTTATPA----- 1178
1134 VGQGTPLVPHEEAAAPGRTTATPAGASLEWSQARGLLFPAPQPERLLPGQENSVQSS 1193

1179 -----GACGRQHLEPTGIDMRGPQADCAVAIGRPLGEVVTLRVLESSLN 1224
1194 YVLSFLSSCCRRGACGRQHLEPTGIDMRGPQADCAVAIGRPLGEVVTLRVLESSLN 1253
1225 CSAGDMLLWGLRUTWRKMKRKLDMTFSSNTNLTIVRQCRPGGVLLRYGSQLAPETF 1284
1254 CSAGDMLLWGLRUTWRKMKRKLDMTFSSNTNLTIVRQCRPGGVLLRYGSQLAPETF 1313
1285 YRECDMOLGPGWGEIYVSPSLSPATSNAGGCELFINVAPHARIAHALATNMGAGTEGANA 1344
1314 YRECDMOLGPGWGEIYVSPSLSPATSNAGGCELFINVAPHARIAHALATNMGAGTEGANA 1373
1345 SYLIRDTLSLRTAFHGOQVLYWESSQAEMEFSEGFKAQASLRGQYWTLSQWVPEM 1404
1374 SYLIRDTLSLRTAFHGOQVLYWESSQAEMEFSEGFKAQASLRGQYWTLSQWVPEM 1433
1405 QDPQSWKXKEGT 1416
1434 QDPQSWKXKEGT 1445

RESULT 5
AAE24449
ID AAE24449 standard; protein; 1353 AA.
XX AAE24449;
AC AAE24449;
XX
DT 04-OCT-2002 (first entry)
XX
XX Human Von Willebrand factor-cleaving protease fragment #2.
DE
XX
XX Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;
KW transgenic animal; immunisation; thromboembolic disease; preclampsia;
KW thrombotic thrombocytopenic purpura; TTP; Henoch-Schonlein purpura;
KW thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;
KW transgenic; anticoagulant.
XX
OS Homo sapiens.
XX
PN WO200242441-A2.
XX
PD 30-MAY-2002.
XX
PF 20-NOV-2001; 2001WO-EP013391.
XX
PR 22-NOV-2000; 2000US-00721254.
PR 12-APR-2001; 2001US-00833328.
XX
XX (BAXT) BAXTER AG.
XX
XX Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
PI Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
PI Zimmermann K, Voelkel D;
XX
XX WPI; 2002-479950/51.
XX
PT Novel isolated or substantially purified Von Willebrand factor-cleaving
protease, useful for producing preparation for therapy of thrombosis and
thromboembolic disease such as thrombotic thrombocytopenic purpura.
XX
PS Claim 1; Page 64-68; 93pp; English.
XX
XX The invention relates to an isolated or substantially pure Von Willebrand
factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
purifying vWF which involves providing vWF-cp as a ligand, contacting a
solution comprising vWF with the polypeptide ligand under conditions
where vWF is bound to the ligand and recovering from the ligand purified
vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
which involves immunising an animal with vWF-cp and isolating the anti-
vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
producing a preparation of prophylaxis and therapy of thrombosis and
thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
Henoch-Schonlein purpura, preclampsia, neonatal thrombocytopenia or

CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing
CC plasmatic or recombinantly produced vWF. The invention is useful for
CC construction expression systems and generating transgenic animals which
CC express the polypeptide in vivo. The present sequence is human vWF-cp
CC fragment
XX
SQ Sequence 1353 AA;

Query Match		92.8%;	Score 7190;	DB 5;	Length 1353;
Best Local Similarity		97.5%;	Pred. No. 0;		
Matches 1319;		Conservative 0;	Mismatches 0;	Indels 34;	Gaps 1;
Qy	98	AAGGILHLELLNAVGPDPVFOAHQEDTERRYVLTNLNIGAEILLRDPISLGAQFRVHLVGMVIL	157		
Db	1	AAGGILHLELLNAVGPDPVFOAHQEDTERRYVLTNLNIGAEILLRDPISLGAQFRVHLVGMVIL	60		
Qy	158	TEPEGAPNTANLTSSLLSVCHSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG	217		
Db	61	TEPEGAPNTANLTSSLLSVCHSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG	120		
Qy	218	VTQLGGACSPWTSCLITDGFGLGVTTIAHEIGHSGFLEHGDGAPSGCGSPGSHVNASDGA	277		
Db	121	VTQLGGACSPWTSCLITDGFGLGVTTIAHEIGHSGFLEHGDGAPSGCGSPGSHVNASDGA	180		
Qy	278	APRAGLAWSPCSRQLLSLAGRARCVMDPPRPPQGSAGHPDPAQPLGYYSANEQCRAVA	337		
Db	181	APRAGLAWSPCSRQLLSLAGRARCVMDPPRPPQGSAGHPDPAQPLGYYSANEQCRAVA	240		
Qy	338	FGPKAVACTFARHLDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCGRCSRSL	397		
Db	241	FGPKAVACTFARHLDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCGRCSRSL	300		
Qy	398	VELTPTAAVHGRWSSWGPSPRCSFSCSGGVVTRRQCNPRPAPFGGRACVADLOAECN	457		
Db	301	VELTPTAAVHGRWSSWGPSPRCSFSCSGGVVTRRQCNPRPAPFGGRACVADLOAECN	360		
Qy	458	TOACEKTQLEFMSQOQARTQQLSPGSGAFYHMGAAVPHSGDALCRMCMRAIGESF	517		
Db	361	TOACEKTQLEFMSQOQARTQQLSPGSGAFYHMGAAVPHSGDALCRMCMRAIGESF	420		
Qy	518	IMKRGDSFLDGTGRCMPGREDGTLSLCVSGSCRTFGCDGRMDSQVWDRQVCGGDNST	577		
Db	421	IMKRGDSFLDGTGRCMPGREDGTLSLCVSGSCRTFGCDGRMDSQVWDRQVCGGDNST	480		
Qy	578	CSPRKGSFTAGRAREYVFTVTNLTSVYIANHRPLFTHLAVRIGRYVWAGKMSISPN	637		
Db	481	CSPRKGSFTAGRAREYVFTVTNLTSVYIANHRPLFTHLAVRIGRYVWAGKMSISPN	540		
Qy	638	TTYPSSLLEDGRVRYVALTEDRLPRLEIRIWGPLEADADIQVYRRYGEYGNLTRPDIT	697		
Db	541	TTYPSSLLEDGRVRYVALTEDRLPRLEIRIWGPLEADADIQVYRRYGEYGNLTRPDIT	600		
Qy	698	FTYFQPKPROAWYAAVVRGSPCSVSCGAGLRWVNYSCLDQARKELVTVQCQSQQPRAWP	757		
Db	601	FTYFQPKPROAWYAAVVRGSPCSVSCGAGLRWVNYSCLDQARKELVTVQCQSQQPRAWP	660		
Qy	758	EACVLPFCPPYAWGDFGFCPSASCGGGLRERFVRCVEAQGSLLKTLPPARCAGAAQQA	817		
Db	661	EACVLPFCPPYAWGDFGFCPSASCGGGLRERFVRCVEAQGSLLKTLPPARCAGAAQQA	720		
Qy	818	ALBETCPQPCARWEVSEPSCTSGAGLAGLALENETCVPGADGLEAPVTEGSGVDEKLP	877		
Db	721	ALBETCPQPCARWEVSEPSCTSGAGLAGLALENETCVPGADGLEAPVTEGSGVDEKLP	780		
Qy	878	APPEPCVGMSCPPGNGHLDATSAKAPSPWGSIRITGAQAHHWTTPAAGSCSVSCGRGLME	937		
Db	781	APPEPCVGMSCPPGNGHLDATSAKAPSPWGSIRITGAQAHHWTTPAAGSCSVSCGRGLME	840		
Qy	938	LRLFCMDSALRVPVQBELCGLASKPGSRREVCQAVPCPARWQYKLAACSVSCGRGVVRI	997		
Db	841	LRLFCMDSALRVPVQBELCGLASKPGSRREVCQAVPCPARWQYKLAACSVSCGRGVVRI	900		
Qy	998	LYCARAHGEDDGBEILLDTQCQGLPRPEPQACSLFPCPPRWKYNLSLGPCSASCGLGTAR	1057		

Db	901	LYCARAHGEDDGBEILLDTQCQGLPRPEPQACSLFPCPPRWKYNLSLGPCSASCGLGTAR	960		
Qy	1058	RSVACVQLDQGGQDVEDEAACAAALVRPEASVPCLIADCTTAYWHVGTWMECSVSCGDGIOR	1117		
Db	961	RSVACVQLDQGGQDVEDEAACAAALVRPEASVPCLIADCTTAYWHVGTWMECSVSCGDGIOR	1020		
Qy	1118	RRDTCILGPOAQAPVPADFCOHLPKPVTVRGCVGQGTSPSLVPHEEAAAPGRTTATP	1177		
Db	1021	RRDTCILGPOAQAPVPADFCOHLPKPVTVRGCVGQGTSPSLVPHEEAAAPGRTTATP	1080		
Qy	1178	AG-----ACGRQHLPTGTIDMRGPGQADCA	1203		
Db	1081	AGASLEWSQARGLLFSFAPQPRELLFPQENSVQSSACGRQHLPTGTIDMRGPGQADCA	1140		
Qy	1204	VAIGRPLGEVVTURVLESSLNCAGDMLLWGRLLTRWKVCRKLLDMTSSKNTNLTIVVROR	1263		
Db	1141	VAIGRPLGEVVTURVLESSLNCAGDMLLWGRLLTRWKVCRKLLDMTSSKNTNLTIVVROR	1200		
Qy	1264	CGRPGGGVLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRFLFNVAPH	1323		
Db	1201	CGRPGGGVLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRFLFNVAPH	1260		
Qy	1324	ARIATHALATNMAGTEGANASYILIRDTLSLRTTAFHGQVLYWESSSSQAEEMFSEGF	1383		
Db	1261	ARIATHALATNMAGTEGANASYILIRDTLSLRTTAFHGQVLYWESSSSQAEEMFSEGF	1320		
Qy	1384	LKAQASLRGQYWTLOSVPWPEMDPQSWKKEGT	1416		
Db	1321	LKAQASLRGQYWTLOSVPWPEMDPQSWKKEGT	1353		

RESULT 6

AAO16617
ID AAO16617 standard; protein; 1353 AA.

XX AAO16617;

XX 15-MAY-2003 (first entry)

XX Human von Willebrand factor (vWF)-cleaving enzyme-related protein #4.

XX Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
XX von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
XX myocardial infarction; cerebral infarction; arteriosclerosis;
XX platelet thrombosis; stenosis.

XX Homo sapiens.

XX WO200289366-A1.

XX 07-NOV-2002.

XX 25-APR-2002; 2002WO-JP004141.

XX 25-APR-2001; 2001JP-00128342.

XX 27-JUL-2001; 2001JP-00227510.

XX 28-SEP-2001; 2001JP-00302977.

XX 25-JAN-2002; 2002JP-00017596.

XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX WPI; 2003-120479/11.

XX N-PSDB; ABT32583.

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
XX supplementary therapy for, thrombotic thrombocytopenic purpura, and for
XX developing drugs for e.g. myocardial infarction and cerebral infarction.

XX Claim 4; Page 82-92; 144pp; Japanese.

XX

PS Claim 4; Page 92-101; 144pp; Japanese.

CC The invention comprises the amino acid and coding sequence of a von

CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of

CC the invention are useful in the diagnosis and treatment of

CC thrombocytopenic purpura, and in developing drugs for myocardial

CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,

CC and stenosis. The present amino acid sequence represents a human von

CC Willebrand factor (vWF)-cleaving enzyme-related protein

XX Sequence 1297 AA;

QY Query Match 91.4%; Score 7081; DB 6; Length 1297;

DB Best Local Similarity 98.3%; Pred. No. 0;

DB Matches 1297; Conservative 0; Mismatches 0; Indels 22; Gaps 1;

QY 98 AAGGTLHLELVAVGPDVQAHQEDTERVYLTNLTGAEILLRDPSLGAQPRVHLVKVIL 157

DB 1 AAGGTLHLELVAVGPDVQAHQEDTERVYLTNLTGAEILLRDPSLGAQPRVHLVKVIL 60

QY 158 TEPEGAPNITANLTSLLSVCGWQINPDDTDPGHADLVLYITRDELPDGNRQVRG 217

DB 61 TEPEGAPNITANLTSLLSVCGWQINPDDTDPGHADLVLYITRDELPDGNRQVRG 120

QY 218 VTQLGACSPWTSCLITDTPGLVGTIAHEIGHSGFLEHDGAPGSCGPGSHVMSDGA 277

DB 121 VTQLGACSPWTSCLITDTPGLVGTIAHEIGHSGFLEHDGAPGSCGPGSHVMSDGA 180

QY 278 APRAGLAWSPCRRQLLSLSAGRACVMPDPPRQPSAGHPDPAQGLYISANEQCRVA 337

DB 181 APRAGLAWSPCRRQLLSLSAGRACVMPDPPRQPSAGHPDPAQGLYISANEQCRVA 240

QY 338 FGPKAVACTFAHEHLDLMOALSCHTDPLODSCSRLVPLLDGTECGVEKWCGRCSRSL 397

DB 241 FGPKAVACTFAHEHLDLMOALSCHTDPLODSCSRLVPLLDGTECGVEKWCGRCSRSL 300

QY 398 VELTPIAAVHGKWSWGPRSPCSRSCGGVTVTRRRQCNRPAPFGGRACVAGADLQAEKCN 457

DB 301 VELTPIAAVHGKWSWGPRSPCSRSCGGVTVTRRRQCNRPAPFGGRACVAGADLQAEKCN 360

QY 458 TOACEKTLQLEFMSQCCARTDQGLRSPGASFYHGAAPVHSGDLCALCRMCAIGESF 517

DB 361 TOACEKTLQLEFMSQCCARTDQGLRSPGASFYHGAAPVHSGDLCALCRMCAIGESF 420

QY 518 IMKRGDSFLDGTFCMPSGPRDGTSLVCSGSCRTFCGCDRMDSQVWDRQVCGDNST 577

DB 421 IMKRGDSFLDGTFCMPSGPRDGTSLVCSGSCRTFCGCDRMDSQVWDRQVCGDNST 480

QY 578 CSPRKSFTAGRAREVYVTLVTPNLTSVYIANHRPLFTLAVETIGRYVYVAGKMSISPN 637

DB 481 CSPRKSFTAGRAREVYVTLVTPNLTSVYIANHRPLFTLAVETIGRYVYVAGKMSISPN 540

QY 638 TTPYSLLEDGRVEYRVALTEDRLPRLEIRIWGPLEQDADIQVYRRYGEYGNLTRPDIT 697

DB 541 TTPYSLLEDGRVEYRVALTEDRLPRLEIRIWGPLEQDADIQVYRRYGEYGNLTRPDIT 600

QY 698 FTYFQFKPQAWAAVARGPCSVSCGAGLRWVNSCLDQARKELVETVQCQSQPPAMP 757

DB 601 FTYFQFKPQAWAAVARGPCSVSCGAGLRWVNSCLDQARKELVETVQCQSQPPAMP 660

QY 758 EACVLEPCPPYVAVGDFGPGCSAGCGGLRBRVRCVEAQSGLLKLTPPARCAGAQPAV 817

DB 661 EACVLEPCPPYVAVGDFGPGCSAGCGGLRBRVRCVEAQSGLLKLTPPARCAGAQPAV 720

QY 818 ALETNCPQPCAPWEVSEPSSTCSAGGAGLALNETCVPGADGLEAPVTEPGSVDEKLP 877

DB 721 ALETNCPQPCAPWEVSEPSSTCSAGGAGLALNETCVPGADGLEAPVTEPGSVDEKLP 780

QY 878 APEPCVGMSCPPGWHLDATSAEKAPSPWGSIRTGAAHVWTFPAAGSCSVSCGRGLME 937

DB 781 APEPCVGMSCPPGWHLDATSAEKAPSPWGSIRTGAAHVWTFPAAGSCSVSCGRGLME 840

QY 938 LRFLCWDLSALRVFVQEEELCGLASKPGSRREVQAVPCFARWQYKLAACSVSCGRGVVRI 997

DB 841 LRFLCWDLSALRVFVQEEELCGLASKPGSRREVQAVPCFARWQYKLAACSVSCGRGVVRI 900

QY 998 LYCARAHGEDDGEELLDTCQGLPRPEQACSLPEPCPPRWKVMYSLGSPCSASCGLTAR 1057

DB 901 LYCARAHGEDDGEELLDTCQGLPRPEQACSLPEPCPPRWKVMYSLGSPCSASCGLTAR 960

QY 1058 RSVACVQLDQGDQVEDEAAALVRPEASVPCLIADCTYRWHVGTWMECSVSCGDGIQR 1117

DB 961 RSVACVQLDQGDQVEDEAAALVRPEASVPCLIADCTYRWHVGTWMECSVSCGDGIQR 1020

QY 1118 RDTCLGPOAQAPVPADFCQHLPKPVTVRGCMAGPCVGOGTFSVLPHEEAAAAPGRTTATP 1177

DB 1021 RDTCLGPOAQAPVPADFCQHLPKPVTVRGCMAGPCVGO----- 1059

QY 1178 AGACGRQHLPEPTGIDMRPGQADCAVAIGRPLGEVVTLRVLESSLNCAGDMLLLWGRL 1237

DB 1060 -GACGRQHLPEPTGIDMRPGQADCAVAIGRPLGEVVTLRVLESSLNCAGDMLLLWGRL 1118

QY 1238 TWRKMKRLLDMTFFSKTNTLVVRQCGRPGGVLLRYGSQLAPETTFYRECQMQLFGPMW 1297

DB 1119 TWRKMKRLLDMTFFSKTNTLVVRQCGRPGGVLLRYGSQLAPETTFYRECQMQLFGPMW 1178

QY 1298 EIVSPSLSPATSNAGCRLFINVAPHARIAHALATNMGAGTEGANASYIILRDTHSLRT 1357

DB 1179 EIVSPSLSPATSNAGCRLFINVAPHARIAHALATNMGAGTEGANASYIILRDTHSLRT 1238

QY 1358 TAFHGOQVLYWSESSQAEMERSEGLKQAASLRGQVYTLQSWVPEMDQPSWKKEGT 1416

DB 1239 TAFHGOQVLYWSESSQAEMERSEGLKQAASLRGQVYTLQSWVPEMDQPSWKKEGT 1297

RESULT 8

AAO16619

ID AAO16619 standard; protein; 1378 AA.

XX

AC AAO16619;

XX

DT 15-MAY-2003 (first entry)

XX

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #6.

XX

KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

KW platelet thrombosis; stenosis.

XX

OS Homo sapiens.

XX

PN WO200288366-A1.

XX

PD 07-NOV-2002.

XX

PF 25-APR-2002; 2002WO-JP004141.

XX

PR 25-APR-2001; 2001JP-00128342.

PR 27-JUL-2001; 2001JP-00227510.

PR 28-SEP-2001; 2001JP-00302977.

PR 25-JAN-2002; 2002JP-00017596.

XX

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

PI WPI; 2003-120479/11.

XX N-PSDB; ABT32585.

DR

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and

PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for

PT developing drugs for e.g. myocardial infarction and cerebral infarction.

XX

PS Claim 4; Page 102-111; 144pp; Japanese.

XX

CC The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein
XX

SQ Sequence 1378 AA;

Query Match 88.5%; Score 6861.5; DB 6; Length 1378;
Best Local Similarity 92.6%; Pred. No. 0;
Matches 1279; Conservative 4; Mismatches 33; Indels 65; Gaps 5;

QY 98 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNINICAEILLRDSLAGQFRVHLVGMVIL 157
DB 1 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNINICAEILLRDSLAGQFRVHLVGMVIL 60
QY 158 TEPEGAPNTANITSLISVCGHSQTNIPDDTDPGHADLVLYITRFDLELPDGNQVRG 217
DB 61 TEPEGAPNTANITSLISVCGHSQTNIPDDTDPGHADLVLYITRFDLELPDGNQVRG 120
QY 218 VTQLGACSPWTSCLITDGFGLGVITIAEIHGSHFGLHDGAPGCGPSGHVWASDGA 277
DB 121 VTQLGACSPWTSCLITDGFGLGVITIAEIHGSHFGLHDGAPGCGPSGHVWASDGA 180
QY 278 APRAGLAWSPCSRRLSLLSAG-----RARCWMDPPR 310
DB 181 APRAGLAWSPCSRRLSLRLTGALRVGPAASSTRVGPAGCAAPLLQRAVPRGLR 240
QY 311 PPGSGAGHPDAPGLYSANEQCVAF-GPKAVACTFAREHLDQCALSHCTDPLDQSS 369
DB 241 PQ-GCRHLRQAGAF--ESAGGGGLGLVRSIRITQLTSPQTCMDQCALSCTDPLDQSS 297
QY 370 CSRLLVPLLDGTGCEVKEKSKGRCSIVELTPTAAVHGRWSSWGPRSPCSRSGGGVVT 429
DB 298 CSRLLVPLLDGTGCEVKEKSKGRCSIVELTPTAAVHGRWSSWGPRSPCSRSGGGVVT 357
QY 430 RRRQCNPRPAFGRACVGADLQAEQNTQACEKTLQEFMSQCCARTDQGPLRSSPGGAS 489
DB 358 RRRQCNPRPAFGRACVGADLQAEQNTQACEKTLQEFMSQCCARTDQGPLRSSPGGAS 417
QY 490 FYHWGAAPVHSGDALCRHMCRAIGESFIMKRGDSFLDGTGCMPSGPRDGTLSLCVSGS 549
DB 418 FYHWGAAPVHSGDALCRHMCRAIGESFIMKRGDSFLDGTGCMPSGPRDGTLSLCVSGS 477
QY 550 CRTFGCDGRWDSQVWDRQVCGGDNSTCSRKGSFTAGRAREYVTLTVPNLTSVYIA 609
DB 478 CRTFGCDGRWDSQVWDRQVCGGDNSTCSRKGSFTAGRAREYVTLTVPNLTSVYIA 537
QY 610 NHRPLFTHLAVRIGRYYVAGKMSISPNNTTYPSSLLEDGRVEYRVALTDRPLRLEIRW 669
DB 538 NHRPLFTHLAVRIGRYYVAGKMSISPNNTTYPSSLLEDGRVEYRVALTDRPLRLEIRW 597
QY 670 GPIQEDADIQVYRYGSEYGNLTPDITFTYFQPKRQAWVAAVVRGCSVSCGAGLRWV 729
DB 598 GPIQEDADIQVYRYGSEYGNLTPDITFTYFQPKRQAWVAAVVRGCSVSCGAGLRWV 657
QY 730 NYSCLDQARKELVETVQCQSQQPPAWPEACVLEPCPPYAWVGDGFCPCASCGGGLRERP 789
DB 658 NYSCLDQARKELVETVQCQSQQPPAWPEACVLEPCPPYAWVGDGFCPCASCGGGLRERP 717
QY 790 VRCVEAGSLLKTLPPARCAGACQAPVALETCNPQCFARWEVSEPSCTSAGAGLAL 849
DB 718 VRCVEAGSLLKTLPPARCAGACQAPVALETCNPQCFARWEVSEPSCTSAGAGLAL 777
QY 850 ENETCVPGADGLEAPVTEGPGSVDEKLPAPCEVGMSCPPGWHGLDTSAGKAPSPWGS 909
DB 778 ENETCVPGADGLEAPVTEGPGSVDEKLPAPCEVGMSCPPGWHGLDTSAGKAPSPWGS 837
QY 910 IRTGAQAAHVWTPAAGSCSVSCGREGLMELRFLCMDLSALRVPVQEEELCGLASKPGSRREVC 969
DB 838 IRTGAQAAHVWTPAAGSCSVSCGREGLMELRFLCMDLSALRVPVQEEELCGLASKPGSRREVC 897

QY 970 QAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTCCQSLPRPEQEA 1029
DB 898 QAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTCCQSLPRPEQEA 957
QY 1030 CSLEPCPPRWKMSLGPSCASCOLGTARRSVACVQLDQGDQVDEDAACALVPEASVP 1089
DB 958 CSLEPCPPRWKMSLGPSCASCOLGTARRSVACVQLDQGDQVDEDAACALVPEASVP 1017
QY 1090 CLIACTYRWHVGTWMECSVSCGDGIQRRDTCGLGPOAQAPVPADFCQHLKPKVTVRGW 1149
DB 1018 CLIACTYRWHVGTWMECSVSCGDGIQRRDTCGLGPOAQAPVPADFCQHLKPKVTVRGW 1077
QY 1150 AGPCVQGGTPSLVPEHEAAAPGRTTATPAG----- 1179
DB 1078 AGPCVQGGTPSLVPEHEAAAPGRTTATPAGASLEWSQARGLLFSPAPQPRELLPGQENS 1137
QY 1180 ----ACGRQHLEPTGTIDMRGPQOADCVAIGRPLGEVVTLRVLESSLNCAGDMLLLWG 1235
DB 1138 VQSSACGRQHLEPTGTIDMRGPQOADCVAIGRPLGEVVTLRVLESSLNCAGDMLLLWG 1197
QY 1236 RLTRWRCRKLDMTFTSSKTNTLVQRQCRPGGVLLRVGSQLAPETFYRECDMLRGP 1295
DB 1198 RLTRWRCRKLDMTFTSSKTNTLVQRQCRPGGVLLRVGSQLAPETFYRECDMLRGP 1257
QY 1296 WGEIVSPSPATSNAGGCELPINVAPHARIAIHALATNMGAGTEGANASVILIRDTHSL 1355
DB 1258 WGEIVSPSPATSNAGGCELPINVAPHARIAIHALATNMGAGTEGANASVILIRDTHSL 1317
QY 1356 RTTAFHGQVLYWSESSQAESEFSEGLKQAQSLRGQYWTLSQWVPEMQDPQSWKKEG 1415
DB 1318 RTTAFHGQVLYWSESSQAESEFSEGLKQAQSLRGQYWTLSQWVPEMQDPQSWKKEG 1377
QY 1416 T 1416
DB 1378 T 1378
RESULT 9
AAO16620
ID AAO16620 standard; protein; 1323 AA.
XX AAO16620;
XX
DT 15-MAY-2003 (first entry)
XX
DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #7.
XX
KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
KW myocardial infarction; cerebral infarction; arteriosclerosis;
KW platelet thrombosis; stenosis.
XX
OS Homo sapiens.
XX
PN WO200288366-A1.
XX
PD 07-NOV-2002.
XX
PF 25-APR-2002; 2002WO-JP0004141.
XX
PR 25-APR-2001; 2001JP-00128342.
PR 27-JUL-2001; 2001JP-00227510.
PR 28-SEP-2001; 2001JP-00302977.
PR 23-JAN-2002; 2002JP-00017596.
XX
PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
XX
PI Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
XX
DR WPI; 2003-120479/11.
DR N-PSDB; ABT32586.
XX

PT von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
PT developing drugs for e.g. myocardial infarction and cerebral infarction.

XX PS PS
XX Claim 4; Page 112-121; 144pp; Japanese.

CC The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombotic thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein

XX SQ Sequence 1323 AA;

Query Match 87.1%; Score 6752.5; DB 6; Length 1323;
Best Local Similarity 93.3%; Pred. No. 0;
Matches 1257; Conservative 4; Mismatches 33; Indels 53; Gaps 5;

Qy 98 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDPGLAQFRVHLVKNVIL 157
Db 1 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDPGLAQFRVHLVKNVIL 60
Qy 158 TEPEGAPNITANLTSSLLSVCGWSQINPDDTDPGHADLVLYITRFDELPLDGNQVRG 217
Db 61 TEPEGAPNITANLTSSLLSVCGWSQINPDDTDPGHADLVLYITRFDELPLDGNQVRG 120
Qy 218 VTOLGACSPWTSCLTETDGLVGTIAHEIGHSGFLEHGDGPGSGGSHVNASDGA 277
Db 121 VTOLGACSPWTSCLTETDGLVGTIAHEIGHSGFLEHGDGPGSGGSHVNASDGA 180
Qy 278 APRAGLAWSPCRRQLLSLSAG-----PARCVMDDPPR 310
Db 181 APRAGLAWSPCRRQLLSLTALRGVPAASSTRVTRGAPAGCAAMPILQQRVPRGLR 240
Qy 311 PPGSAGHPDPAQGLYYSANECRVAF-GPKAVACTFAREHLMQALSCHTDLPLDQSS 369
Db 241 PQ-GCRLLHLRQAGAPG--ESAGGGLGLAVRLSRLTQLTSPQTCMDMCQALSCHTDLPLDQSS 297
Qy 370 CSRLVPLLDGTGECVGEKWKCKRCSRLVELPTIAAVHGRWSSNGPSPCSRS CGGGVYT 429
Db 298 CSRLVPLLDGTGECVGEKWKCKRCSRLVELPTIAAVHGRWSSNGPSPCSRS CGGGVYT 357
Qy 430 RRRQCNPRPAGFCACVADLAEMCNTQACEKTQLEFMSQCCARTDQGLRSPGGAS 489
Db 358 RRRQCNPRPAGFCACVADLAEMCNTQACEKTQLEFMSQCCARTDQGLRSPGGAS 417
Qy 490 FYHWGAAPHSQGDALCRHMCRAIGESFIMKRGDSFLDGTGRCMPSPGPDGTLSLCVSGS 549
Db 418 FYHWGAAPHSQGDALCRHMCRAIGESFIMKRGDSFLDGTGRCMPSPGPDGTLSLCVSGS 477
Qy 550 CRTFGCDGRMDSQVWDRQVCGGDNSTCSPRKGSFTAGRAREYVFTLVTPLNLTSVYIA 609
Db 478 CRTFGCDGRMDSQVWDRQVCGGDNSTCSPRKGSFTAGRAREYVFTLVTPLNLTSVYIA 537
Qy 610 NHRPLTHLAVRIGGRVWAGKMSISNTYTPSLLEDGRVYRVALTEDRLPRLEEIRIM 669
Db 538 NHRPLTHLAVRIGGRVWAGKMSISNTYTPSLLEDGRVYRVALTEDRLPRLEEIRIM 597
Qy 670 GLQEDADIQVRYRYGEYGNLTPDITFTYFQPKPQAWVWAAVRGPCVSCGAGLRWV 729
Db 598 GLQEDADIQVRYRYGEYGNLTPDITFTYFQPKPQAWVWAAVRGPCVSCGAGLRWV 657
Qy 730 NYSCLDQARKELVETVQCQSQSQPPAMPEACVLEPCPPYAVAGDFGPCSASCGGLRERP 789
Db 658 NYSCLDQARKELVETVQCQSQSQPPAMPEACVLEPCPPYAVAGDFGPCSASCGGLRERP 717
Qy 790 VRCVEAQGSLKLTLPARCRAGAQQPAVALETCNPPQPCPARWEVSEPSCTSAGAGLAL 849
Db 718 VRCVEAQGSLKLTLPARCRAGAQQPAVALETCNPPQPCPARWEVSEPSCTSAGAGLAL 777
Qy 850 ENETCVPGADGLEAPVTEGPGSVDEKLPAPPCVGMSCPPGWHGLDATSAGEKAPSPWGS 909

Db 778 ENETCVPGADGLEAPVTEGPGSVDEKLPAPPCVGMSCPPGWHGLDATSAGEKAPSPWGS 837
Qy 910 IRTGAQAHHVWTPAAGSCSVSCGRGLMELRFLCWDLSALRVPVQBELCGLASKPGSRREVC 969
Db 838 IRTGAQAHHVWTPAAGSCSVSCGRGLMELRFLCWDLSALRVPVQBELCGLASKPGSRREVC 897
Qy 970 QAVPCPARQWYKLAACSVSCGRGVVRILYCARAHGEDDGEIILLDTQCQGLPRPEPQEA 1029
Db 898 QAVPCPARQWYKLAACSVSCGRGVVRILYCARAHGEDDGEIILLDTQCQGLPRPEPQEA 957
Qy 1030 CSLEPCPPKWKMSLQPCASCSGLGTARRSVACVQLDQGDQVEVDEAACAALVREASVP 1089
Db 958 CSLEPCPPKWKMSLQPCASCSGLGTARRSVACVQLDQGDQVEVDEAACAALVREASVP 1017
Qy 1090 CLIACTYERHVTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFCQHLKPKVTVRGW 1149
Db 1018 CLIACTYERHVTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFCQHLKPKVTVRGW 1077
Qy 1150 AGPCVQGTPLVPHBEAAAPGRTTATPAGACGROHLEPTGTIDMRGPGQADCAVAIGRP 1209
Db 1078 AGPCVQGTPLVPHBEAAAPGRTTATPAGACGROHLEPTGTIDMRGPGQADCAVAIGRP 1115
Qy 1210 LGVVTLRVLESSLNCSAGDMLLWGLRTWRKCKRLDMTFSSKTNLTVVQRQCRPGG 1269
Db 1116 LGVVTLRVLESSLNCSAGDMLLWGLRTWRKCKRLDMTFSSKTNLTVVQRQCRPGG 1175
Qy 1270 GYLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRFLFINVAPHARIAH 1329
Db 1176 GYLLRYGSQLAPETFYRECDMQLFGPWGEIVSPSLSPATSNAGGCRFLFINVAPHARIAH 1235
Qy 1330 ALATNMGAGTEGANASYILIRDTHTSLRTTAFHGQQVLYWSESSQAEMFSEGFLLKQAAS 1389
Db 1236 ALATNMGAGTEGANASYILIRDTHTSLRTTAFHGQQVLYWSESSQAEMFSEGFLLKQAAS 1295
Qy 1390 LRQYWTLOSVPWEMQDPSWKKEGT 1416
Db 1296 LRQYWTLOSVPWEMQDPSWKKEGT 1322

RESULT 10
AAU79217
ID AAU79217 standard; protein; 933 AA.
XX
AC AAU79217;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human ADAM-TS-like protein.
XX
KW Human: ADAM-TS-like protein; cardiovascular disorder; angina;
KW vascular system; congestive heart failure; myocardial infarction;
KW ischaemic heart disease; arrhythmia; hypertensive vascular disease;
KW secondary arterial hypertension; peripheral vascular disease; embolism;
KW chronic peripheral arterial occlusive disease; acute arterial thrombosis;
KW inflammatory vascular disorder; chronic obstructive pulmonary disease;
KW liver disorder.
XX
OS Homo sapiens.
XX
PN WO200226999-A2.
XX
PD 04-APR-2002.
XX
PF 26-SEP-2001; 2001WO-EF011124.
XX
PR 28-SEP-2000; 2000US-0235881P.
PR 25-JUL-2001; 2001US-0307393P.
XX
PA (FARB) BAYER AG.
XX
PI Xiao Y;
XX

PT New anti-angiogenic intestinal polypeptides, zdnits polypeptides, which
PT are members of disintegrin proteases, for modulating extracellular matrix
PT interaction, tumor suppression and wound healing.
XX
XX Claim 2; Page 84-88; 92pp; English.

XX The present sequence represents a human zdnits polypeptide. The zdnits
CC polypeptide is an anti-angiogenic intestinal polypeptide. Zdnits is used
CC for modulating extracellular matrix interactions. Zdnits polypeptide is
CC useful as a tool for identifying new family members of polypeptides.
CC Zdnits polynucleotides are useful as probes or primers to clone 5' non-
CC coding regions of zdnits gene. Zdnits polypeptides are used for tumor
CC suppression, gamete maturation, immunologic recognition, and growth and
CC differentiation either working in isolation or in conjunction with other
CC molecules in colon, small intestine, fetal lung, testis and B-cells.
CC zdnits polypeptides are also useful for promoting wound healing, in the
CC treatment of disorders associated with recovery after gastrointestinal
CC irradiation, chemotherapy or antibody use, as anti-infectives, and
CC extracellular matrix repair and remodeling. The polypeptides are also
CC useful for modulating proteolysis, apoptosis, angiogenesis, infection,
CC cell adhesion, cell fusion and signalling. The polypeptides are also
CC useful for treating tumor formation, Crohn's disease, inflammatory bowel
CC disease, food poisoning, melanoma, degenerative diseases, disorders
CC related to immunity, inflammation, fertility, gamete maturation,
CC immunology, trauma and epithelial disorders

XX SQ Sequence 1120 AA;

Query Match 45.8%; Score 3553.5; DB 4; Length 1120;
Best Local Similarity 54.2%; Pred. No. 2.9e-258;
Matches 766; Conservative 21; Mismatches 108; Indels 517; Gaps 25;

QY 6 KAPSHAPLLGLALLRMHQRHPRPPLCVAGILACGLLGCWGPSPHFOQSCICLQALEPQ 67
DB 15 KAPSHAPLLGLALLRMHQRHPRPPLCVAGILACGLLGCWGPSPHFOQSCICLQALEPQ 74
QY 68 AVSSYLSGAPLKGRRPSPGFQRRQRRAAGGILHLELLVAVGPDVFOAHQEDTERYV 127
DB 75 AVSSYLSGAPLKGRRPSPGFQRRQRRAAGGILHLELLVAVGPDVFOAHQEDTERYV 134
QY 128 LTNLNIAGELLRDPSCAGPRVHLVKVILTEPEGAPNITANLTSLLSVCWGSQTINPE 187
DB 135 LTNLNIAGELLRDPSCAGPRVHLVKVILTEPEGAPNITANLTSLLSVCWGSQTINPE 194
QY 188 DTDGPHADLVYITRFDLELPGNRQVRGVTOLGGACSPWSCILITDGTGDLGVITAH 247
DB 195 DTDGPHADLVYITRFDLELPGNRQVRGVTOLGGACSPWSCILITDGTGDLGVITAH 254
QY 248 EIGHSGFLEHGDGAPGCGSPGHVMSADGAAPRAGLAWSPCSRRLQLS-LLSAGRAR 303
DB 255 EIGHSGFLEHGDGAPGCGSPGHVMSADGAAPRAGLAWSPCSRRLQLSAGPGALRVGTR 314
QY 304 CYWDPFRPQPGSAGHPDAP-GLYYSANQCRVAFGPKAVACTFAREHLD-CCQALS 359
DB 315 GL- -KPGFRGAPAGWRSGLYYSANQCHVAFGPPG- - - - -CRHLRQAGAPCQALS 363
QY 360 CHTDPLDQSSCSRLVPLLDGTGCGVEKWCGRCSRSLVELTPIAAVHGRWSSGPRSPC 419
DB 364 CHTDPLDQSSCSRLVPLLDGTGCGVEK- - - - -VHGRWSSGPRSPC 405
QY 420 SSSCGGVVTRRQCN- - - - -PRPAGRACVGDALQAEKMCNTOACEKTLQEFWSQ 472
DB 406 SSSCG- - - - -CGHQEAQQVQPPQGLFCQGMKLGCGWGTCPSCSVQDTLFHS- - 455
QY 473 CARTDQQLRSPGGASFYHMGAAVPHSQDLCRHMCRALGESFIMKRGD-SFLDGTGRC 531
DB 456 - - - - -ALP- - -GDALCRHMCRALGESKXHEAWETASNGTRC 489
QY 532 MPSPGREDGTLSLCVSGSCRTFGCDGRMSDQVWDRQCQVCGDNSTCSPRKGSTAGRAR 591
DB 490 MPSPGREDGTLSLCVSGSCVRGCDGRMSDQVWDRQCQVCGDNSTCSPRKGSTAGRAR 549
QY 592 EYVTELTVPNTLTSVYIANHRPLFTHLAVRIGGRYVAVGRKMSISFNTTYPSSLEDGRVEY 651

DB 550 EYVTELTVPNTLTSVYIANHRPLFTHLAVRIGGRYVAVGRKMSISFNTTYPSSLEDGRVEY 609
QY 652 RVALTEDRLPRLEETIRWGLQEDADIQVYRRYGEENLTPDITFTYFQPKPQAWW 711
DB 610 - - - - - 609
QY 712 AAVRGPCSVSCAGLRWVNSCLDQARKELVETVQCGSQOPPAWPEACVLEPCPPYWAV 771
DB 610 - - - - -QCVKKIIPGS- - - - -SAYSIN 625
QY 772 GDFGPCSASCGGLRPRVRCVEAQSLLKTLPPARCAGAAQPAVALETNCPQPCPARW 831
DB 626 QDFPVL- - - - -GLRNRTV- - - - -TNTGLR- - - - -HW 646
QY 832 EVSEFSSCTSAGAGLALNETCVPGADGLEAPVTEGPGVDEKLPAPPCVGMSCPGW 891
DB 647 - - - - -PDVGIE- - - - - 652
QY 892 GHLDATSAGEKAPSPWGSIRTAQAHAHVTPAAGSCVSCGRGLMEIRFLCWDMSALRPV 951
DB 653 - - - - -GAGLMEIRFLCWDMSALRPV 672
QY 952 QEELGCLASKPGSRREVCOAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGE 1011
DB 673 QEELGCLASKPGSRREVCOAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGE 732
QY 1012 ILDTQCQGLPRPEQACSLPCPRWKVMSGLPCSCGLGTARRSVACVQLDQGDV 1071
DB 733 ILDTQCQGLPRPEQACSLPCPRWKVMSGLPCSCGLGTARRSVACVQLDQGDV 792
QY 1072 EYDEAACAALVPEASVPCLIADCTYRWHVGTWMEVSCVCGDGIQRRDTCIGPOAQP 1131
DB 793 EYDEAACAALVAAR- - - - - 806
QY 1132 PADFCOHLKPKVTVRCWAGPCVCGTSPSLVPHEEAAAPGTTATPAGACGRHLEPTGT 1191
DB 807 - - - - -GOLSPVSLP- - - - -TAPTAGM- - - - -LAPGWR 828
QY 1192 IDMRGQADCAVIGRPLGEVTVLRVLESSLNC-AGDMLLLWGLTWKRCRLDMT 1250
DB 829 RGSAG- - - - -SLPCRFGLDMLLLWGLTWKRCRLDMT 864
QY 1251 FGSKNTLVVRCGRPGGVLRLYSGQLAPETFYRECQMLFGPWGEIVSPSLSPATSN 1310
DB 865 FSKNTLV- - - - - 873
QY 1311 AGGCRLFINVAPHARIAIHALATNMAGTEGANASYILIRDTSLRTTAFH-GQVLYWE 1369
DB 874 - - - - -IRDTSLRTTAFHRAAGALTGS 895
QY 1370 SSSQAEWEPSEGLKQAASLRQYWTLSQSW 1401
DB 896 QRAARLMEFSEGLKQAASLRQYWTLSQSW 927

RESULT 12

ABG74113

ID ABG74113 standard; protein; 1120 AA.

XX AC ABG74113;

XX AC ABG74113;

DT 01-MAY-2003 (first entry)

XX Human mature zdnits5.

DE Human; zdnits5; ADAM; a disintegrin and metalloprotease; detergent;

KW host defense; tumour; extracellular matrix repair; proteolysis;

KW apoptosis; angiogenesis; bacterial infection; Crohn's disease;

KW inflammatory bowel disease; food poisoning; melanoma;

KW degenerative disease; chromosome 9q34; wound healing; TSP-1;

KW Thrombospondin domain.

XX

OS	Homo sapiens.
XX	
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 474
FT	/label= UNKNOWN
FT	/note= "Encoded by TTN"
FT	1118..1119
FT	/note= "Encoded by CTGTGAAT"
PN	US2002142439-A1.
XX	
XX	
PD	03-OCT-2002.
XX	
PB	09-FEB-2001; 2001US-00781080.
XX	
PR	10-FEB-2000; 2000US-0181511P.
XX	{HOLL} HOLLOWAY J L.
PA	{SHEP} SHEPPARD P O.
PA	{YAMA} YAMAMOTO G.
XX	
PI	Holloway JL, Sheppard PO, Yamamoto G;
XX	
PI	WPI; 2003-1741136/17.
XX	
PS	N-PSDB; ABX15947; ABX15948.
DR	
XX	New zdint5 polypeptides, useful for diagnosing, preventing or treating tumor formation, Crohn's disease, inflammatory bowel disease, food poisoning, melanoma, degenerative disease, bacterial infection, or for wound healing.
XX	Claim 2; Page 32-34; 37pp; English.
CC	The invention relates to an isolated polypeptide, zdint5, (a disintegrin and metalloprotease (ADAM), the gene for which is located on human chromosome 9q34) comprising the metalloprotease domain, thrombospondin-1 domains (TSP-1) and mature protein. Also included are an isolated polynucleotide encoding a fusion protein comprising a first and a second polypeptide segments (where the first polypeptide segment comprises the polypeptide with the sequence of S2, and the second polypeptide comprises one or more (TSP)-like domains, and where the first polynucleotide segment is positioned amino-terminally to the second polynucleotide segment), an expression vector (comprising the following operably linked elements: (a) a transcription promoter; (b) a DNA segment encoding a polypeptide comprising the amino acid sequences listed above; and (c) a transcription terminator), a cultured cell into which the expression vector of has been introduced, modulating extracellular matrix interactions by combining the polypeptides with cells, producing an antibody to the polypeptides and an isolated immunogenic polypeptide comprising at least 30 contiguous amino acids of mature zdint5. The polypeptides and polynucleotides are useful for modulating cellular interactions. The polypeptides and antibodies to zdint5 are useful for in vivo diagnostic or therapeutic applications, such as in identifying or treating tissues or organs that express a corresponding anti-complementary molecule. The polypeptides are useful for diagnosing, preventing or treating tumour formation, Crohn's disease, inflammatory bowel disease, food poisoning, melanoma, degenerative disease, bacterial infections, extracellular matrix repair and remodeling, or for wound healing. The polypeptides are also useful for the study of host defence, extracellular matrix repair, proteolysis, apoptosis, angiogenesis and as a detergent. The antibodies may be used for tagging cells that express zdint5, for isolating zdint5 by affinity purification, or for diagnostic assays for determining circulating levels of zdint5 polypeptides. The present sequence represents mature zdint5
XX	
SQ	Sequence 1120 AA;
Query Match	45.8%; Score 3553.5; DB 6; Length 1120;
Best Local Similarity	54.2%; Pred No. 2.9e-258;
Matches	766; Conservative 21; Mismatches 108; Indels 517; Gaps 25;
8 KAPSHSAPLLGLALRMHQHPRARCPPLCVAGILACGFLLCGWGPHSFQCSCLOALEPQ 67	
15 KAPSHSAPLLGLALRMHQHPRARCPPLCVAGILACGFLLCGWGPHSFQCSCLOALEPQ 74	
68 AVSSYLSPGAPLKGRPPSPGFQRQORRAAAGIILELLVAVGPVFAHQEDTERIV 127	
75 AVSSYLSPGAPLKGRPPSPGFQRQORRAAAGIILELLVAVGPVFAHQEDTERIV 134	
128 LTNLNTGABELLRDPSLGAQFRVHLVKMVIITBEGAPNITANLTSSLISVCWSQTINPE 187	
135 LTNLNIGABELLRDPSLGAQFRVHLVKMVIITPCQAFNITANLTSSLISVCWSQTINPE 194	
188 DTDDPGHADLVLYITRFDLELPDGNRVGVTVQLGGACSPWSCIITDTDFDLGVTAH 247	
195 DTDDPGHADLVLYITRFDLELPDGNRVGVTVQLGGACSPWSCIITDTDFDLGVTAH 254	
248 EIGHSEGLEHDGAPSGCGSPGHVMASDGAAAPRAGLAWSPCSRRLQLLSAGPGALRVGTR 303	
255 EIGHSEGLEHDGAPSGCGSPGHVMASDGAAAPRAGLAWSPCSRRLQLLSAGPGALRVGTR 314	
304 CVWDPRPOPGSAGHPDDAQF-GLYTSANEQCRVAFGPKAVACTFAREHLDL---CQALS 359	
315 GL-----KPGFRGAPAGWESLGLIYSANEQCHVAFGPPG-----CRHLHGAGCAQALS 363	
360 CHTDPLDQSSCSRLLVPLLDGTECGVEKWCKSKRCRSLVELTPIAAVHGRWSSWGPRSPC 419	
364 CHTDPLDQSSCSRLLVPLLDGTECGVEK-----VHGRWSSWGPRSPC 405	
420 SRSCGGVVTRRQCNV-----PRPAGRCVAGADLQAEMQNQACEKTQLEFMSQQ 472	
406 SRSCGG-----CGHQEAAVQOPVPQGGLFCQMKGGLGGWGPCSCSVQDTLFHS-- 455	
473 CARTDQOPLRSSPGGASFYHWGAAPHVSQGDALCRHMCRAIGESPIMKRGD-SFLDGTTC 531	
456 -----ALP-----GDALCRHMCRAIGESKHAEAWETASNGTRC 489	
532 MPSPREDGTLCLVSGSCTRFGCDGRMDSOQVNDRCOVCGDNSTCSPRKGSFTAGRAR 591	
490 MPSPREDGTLCLVSGSCTRFGCDGRMDSOQVNDRCOVCGDNSTCSPRKGSFTAGRAR 549	
592 EYTFELTVTNLTSSVYIANHRPLFTHLVRIGRYVWAGKMSISPTNYPSLLEDGRVGY 651	
550 EYTFELTVTNLTSSVYIANHRPLFTHLVRIGRYVWAGKMSISPTNYPSLLEDGRVGY 609	
652 RVALTEDRLPRELBIRIWGLPLEDAIQVYRRYGEYGNLTRPDITFTFYQPKPQAWVW 711	
610 ----- 609	
712 AAVRGPCSVSCGAGLRWNYSCLDOARKELVETVQCQSQQPPAPWEACVLEPCPYMAV 771	
610 -----QCVKKQIPGS-----SAYSILN 625	
772 GDFGPCSASCGGLRERPVRCVEAAGSLTKLTPARCRAQAQPAVALETNCNPQPAPRW 831	
626 QDFPVL-----GLRNRTV-----TNTGLR-----HW 646	
832 EVSEPSSCTSAGGAGLALENETCVPGADGLEAPVTEGPGSDVKLPAPEPCVMGSCP GW 891	
647 -----PDVGIE----- 652	
892 GHLDATSAGEKAPSPWGSIRTGAAAHVTTPAAGSCSVSCGRGLMELRFLCMDSALRVFPV 951	
653 -----GAGLMELRFLCMDSALRVFPV 672	
952 QEELCGLASKPGSRREVCOAVPCPARWQYKLAACSVCGRGVYRILYCARAAGEDDGE 1011	
673 QEELCGLASKPGSRREVCOAVPCPARWQYKLAACSVCGRGVYRILYCARAAGEDDGE 732	
1012 ILLDTQCGLPRPBPQACSLPECPPRWKMVSLGPCSASCGLGTARRSVACVOLDOGQDV 1071	
733 ILLDTQCGLPRPBPQACSLPECPPRWKMVSLGPCSASCGLGTARRSVACVOLDOGQDV 792	
1072 EVDEAACAAALVRAPEASVPECLDIADTYRHWHGVTWMECSVCGDGIORRDTCIGPQAPV 1131	
793 EVDEAACAAALVRAPEASVPECLDIADTYRHWHGVTWMECSVCGDGIORRDTCIGPQAPV 806	

QY 1132 PADFCQHLPKPVTVRGWAGPCVGQGTSLVPHHEAAAPGRTTATPAGACGRQHLEPTGT 1191
Db 807 -----GQLSPVSLP-----TAPTAGM-----LAPGWR 828
QY 1192 IDMRPGQADCAVAGRELGEVVTIVRLVLESSLNCS-AGDMLLLWGLTWKMKCKLLDMT 1250
Db 829 RGSAGP-----SLPCRFGLGDMLLWGLTWKMKCKLLDMT 864
QY 1251 FSSNTNVLVVRQCRPGGVLLRYGSQLAPETFYRECQMLFGPGEIVSPSLSPATSN 1310
Db 865 FSSNTNVLV-----SLPCRFGLGDMLLWGLTWKMKCKLLDMT 873
QY 1311 AGGCLFNVAPHARIAHALATNMGAGTEGANASVILIRDTSHSLTTAFH-CQQVLYWE 1369
Db 874 -----RDTHSLRTTAFHRAAGALTGS 895
QY 1370 SSSQAEMFSEGFLLKAQASLRGQVWTLQSW 1401
Db 896 QRAARLRNEFSEGFLLKAQASLRGQVWTLQSWL 927
RESULT 13
ABR96171
XX AC ABR96171;
XX DT AC ABR96171;
XX TT 15-SEP-2003 (first entry)
XX DE Human NOV17a protein SEQ ID NO:84.
XX KW Human; NOVX; G protein-coupled receptor; cytostatic; cardiovascular;
KW immunosuppressive; anti-HIV; antiasthmatic; antiarteriosclerotic; AIDS;
KW hypotensive; gene therapy; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; aortic stenosis; atrial septal defect; neoplasm;
KW atrioventricular canal defect; pulmonary stenosis; prostate cancer;
KW uterine cancer; graft versus host disease; multiple sclerosis; GPCR;
KW acquired immunodeficiency syndrome; Crohn's disease; bronchial asthma;
KW chromosome mapping; forensic identification.
XX OS Homo sapiens.
XX PN WO200290568-A2.
XX PD 14-NOV-2002.
XX QY 02-MAY-2002; 2002WO-US014341.
XX PF 03-MAY-2001; 2001US-0288935P.
XX PR 07-MAY-2001; 2001US-0289087P.
XX PR 08-MAY-2001; 2001US-0289620P.
XX PR 09-MAY-2001; 2001US-0289621P.
XX PR 09-MAY-2001; 2001US-0289817P.
XX PR 09-MAY-2001; 2001US-0289818P.
XX PR 11-MAY-2001; 2001US-0290194P.
XX PR 14-MAY-2001; 2001US-0290753P.
XX PR 15-MAY-2001; 2001US-0291189P.
XX PR 16-MAY-2001; 2001US-0291243P.
XX PR 18-MAY-2001; 2001US-0292001P.
XX PR 21-MAY-2001; 2001US-0292374P.
XX PR 22-MAY-2001; 2001US-0292587P.
XX PR 23-MAY-2001; 2001US-0293107P.
XX PR 24-MAY-2001; 2001US-0293589P.
XX PR 25-MAY-2001; 2001US-0293747P.
XX PR 29-MAY-2001; 2001US-0294110P.
XX PR 30-MAY-2001; 2001US-0294434P.
XX PR 14-AUG-2001; 2001US-0312192P.
XX PR 17-AUG-2001; 2001US-0313173P.
XX PR 17-AUG-2001; 2001US-0313187P.
XX PR 12-SEP-2001; 2001US-0318728P.
XX PR 12-SEP-2001; 2001US-0318744P.
XX PR 15-NOV-2001; 2001US-0335910P.

PR 28-NOV-2001; 2001US-0333891P.
PR 28-NOV-2001; 2001US-0333942P.
PR 03-JAN-2002; 2002US-0345776P.
PR 04-JAN-2002; 2002US-0345220P.
PR 01-MAY-2002; 2002US-0013607L.
XX (CURA-) CURAGEN CORP.
XX Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI Edinger SR, Ellerman K, Gangolli EA, Gerlach VL, Gorman L;
PI Gunther E, Herrmann JL, Ji W, Lepley DW, Lewin DA, Li L;
PI MacDougall JR, Malyankar JM, Metz PD, Padigaru M, Patrujan M;
PI Peyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG;
PI Smithson G, Spytek KA, Stone DJ, Taupier RJ, Tchernev VT;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong H, Miller CE,
DR WPI; 2003-111987/10.
DR N-PSDB; ACF16980.
XX
XX
PT New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
PT heart defects, aortic stenosis, atrial septal defect, or atrioventricular
PT canal defect.
XX
PS Claim 1; Page 173; 491pp; English.
XX
CC ACF16939 to ACF17000 encode the human G protein-coupled receptor (GPCR)
CC proteins, designated NOVX proteins, given in ABR96130 to ABR96191. The
CC NOVX sequences can have cytostatic, cardiovascular, antiasthmatic,
CC immunosuppressive, anti-HIV (human immunodeficiency virus), hypotensive
CC and antiarteriosclerotic activities, and can be used in gene therapy.
CC NOVX polypeptides can be used for treating a syndrome associated with a
CC human disease such as a pathology associated with the polypeptide. NOVX
CC polypeptides, polynucleotides and antibodies can be used for treating or
CC preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
CC heart defects, aortic stenosis, atrial septal defect, atrioventricular
CC canal defect, pulmonary stenosis, prostate cancer, uterine cancer,
CC neoplasm, graft versus host disease, acquired immunodeficiency syndrome
CC (AIDS), Crohn's disease, multiple sclerosis, or bronchial asthma. The
CC nucleic acid sequences may be used in chromosome mapping, identifying
CC individual from minute biological samples (tissue typing), and in
CC forensic identification of a biological sample. ACF17001 to ACF17117
CC represent PCR primers and probes for the NOVX sequences, which are used
CC in an example from the present invention
XX
XX Sequence 984 AA;

Query Match 30.2%; Score 2340; DB 6; Length 984;
Best Local Similarity 56.0%; Pred. No. 6.6e-167;
Matches 500; Conservative 29; Mismatches 108; Indels 256; Gaps 21;
QY 53 PSHFQSCLOALEPQAVSVYLSFGAPLKGRP-----PSFGFQQR-----QQRQR 97
Db 192 PVDLRGCTHSSQPSVYPRNRP-----RPWSERSLPCLFPVAEVMQKSGWNEVQUR 245
QY 98 AAGGILHLELVAVGPDVFAHQEDTERVYLTNINIGAEILLRDPISLGAQFVHLVKWVIL 157
Db 246 A-----MKRLVAVGPDVFAHQEDTERVYLTNINIGAEILLRDPISLGAQFVHLVKWVIL 299
QY 158 TEPEGAPNTANLTSSLLSVCGWSQTNIPEDDDTPGHADLVLYIT-REFDLEPDGNRQR 216
Db 300 TEPEGAPNTANLTSSLLSVCGWSQTNIPEDDDTPGHADLVLYITRRDLELPDGNRQR 359
QY 217 GVTQLGGACSPWTSCLITEDTGFGLGVITIAEIGH----- 251
Db 360 GVTQLGGACSPWTSCLITEDTGFGLGVITIAEIGH----- 419
QY 252 -----SFGLEHGDGAPSGCGSPGHVMSADGAAPRAGLAWSPC 288
Db 420 SKQWLVLEFLQRSLYSPRQTQVSFGLHGDGAPSGCGSPGHVMSADGAAPRAGLAWSPC 479
QY 289 SRRQLLSLSAGBARCVWDPRPQPSAGHPDPAQGLYLSANECRCRVAFGFKVACTFA 348

Db 480 SRQLLSLL--GRACVNDPPRPQPSAGHPDPAQPLGLYYANSNQCRVAFGPKAVAC--- 534

Qy 349 REHLDMQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCSCRCRSLVELTPIAAVHG 408

Db 535 ---DMQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCSCRCRSLVELTPIAAVHG 590

Qy 409 RWSWGPRSPCSRSCGGVVTTRQCNPRPAPFGGRACVADLQAEVNTQACEKTQLEF 468

Db 591 RWSWGPRSPCSRSCGGVVTTRQCNPRPAPFGGRACVADLQAEVNTQACEKTQLEF 620

Qy 469 MSQCARTDQQLRSPGSGAFYHGAAPHSO-----GDAL-----CRMCR 511

Db 621 ---YRREGAFSGCPGK-----PEVSHSQLFRASSVHACKLGSVLSDVHQCRMCR 669

Qy 512 AIGESFIMKRGDSFLDGTTCMPSPREDGTLSLCVSGSCRTFGCDGRMDSQQVWDRQVC 571

Db 670 AIGESFIMKRGDSFLDGTTCMPSPREDGTLSLCVSGSCRTFGCDGRMDSQQVWDRQVC 729

Qy 572 GDNSTC-----SPRKGSTFA-----GRAREYVTFLT 598

Db 730 GDNSTCHGVGPRSHQDPGTPTSPPPGRATAPILPAGPRQACGTQGMWASRGQVRVT 789

Qy 599 VTENLTSVVIANHRPLFTHLAVRIGRYVVGKMSISPNTTYPSSLLEDGRVEYRVALTED 658

Db 790 TSP-----IPARPICQLPASVGQRGQPOSQVQGWGPGRSRPTSLNS-----IPD 833

Qy 659 RUPLEIRIWGLQEDADI---QVRRYGEYGNLT-RPD--ITFTYFQPKPQAWYA 712

Db 834 SLPSTQ---GGPQMAHSSAISLSLRHGWYQGMVTSFNHLVVASARIKPKQAWYA 890

Qy 713 AVRGPCSVSCGAGLRWVNSCLDQARKELVETVQCQSQQPAPWPEACVLEPCPPYWAVG 772

Db 891 AVRGPCSVSCGAGETW----- 906

Qy 773 DRGPCSASCGGLRERPVRCVEAQGSLKTLPPARCAGAGQAPVALETCPNPQ 825

Db 907 -----GLIHSTACVEAQGSLKTLPPARCAGAGQAPVALETCPNSE 948

RESULT 14

AAU97642

ID AAU97642 standard; protein; 364 AA.

XX AC AAU97642;

XX 27-AUG-2002 (first entry)

DE Human Hsa011374 protein.

XX Aggrecanase; osteopathic; antiarthritic; antiinflammatory; human;

XX Cartilage; osteoarthritis; inflammatory disease; enzyme; Hsa011374.

OS Homo sapiens.

XX WO200233093-A2.

XX 25-APR-2002.

XX 17-OCT-2001; 2001WO-US032458.

XX 18-OCT-2000; 2000US-0241469P.

XX (GEMY) GENETICS INST INC.

XX Racie LA, Twine NC, Agostino MU, Wolfman NM, Morris EA;

XX WPI; 2002-454602/48.

XX N-PSDE; ABK52580.

XX Novel purified aggrecanase polypeptide useful for developing inhibitors

PT and antibodies to the aggrecanase polypeptide, which are useful for

PT treating aggrecanase-associated condition such as osteoarthritis.

PS Disclosure; Page 37-38; 41pp; English.

XX This invention relates to the cDNA and protein sequences of a novel human

CC aggrecanase polypeptide. The protein of the invention may be used to

CC inhibit the proteolytic activity of aggrecanase, or to inhibit the

CC aggrecanase-mediated cleavage of aggrecan in cartilage. The protein of

CC the invention is useful for developing inhibitors of aggrecanase protein.

CC The cDNA sequence encoding the aggrecanase protein of the invention is

CC useful for designing probes for obtaining DNA sequences encoding other

CC aggrecanase molecules. The cDNA sequence is also useful for detecting

CC detecting or diagnosing genetic disorders involving the aggrecanase, or

CC disorders involving cellular, organ or tissue disorders in which

CC aggrecanase is irregularly transcribed or expressed. The DNA sequences

CC may also be useful for preparing vectors for gene therapy applications.

CC An inhibitor of the protein is useful in treating conditions

CC characterised by degradation of articular cartilage, by blocking the

CC enzyme's proteolytic activity. An aggrecanase protein inhibitor and a

CC method for inhibition of its activity are useful for treating various

CC aggrecanase-associated conditions including osteoarthritis and other

CC inflammatory diseases. The present sequence represents the human

XX Hsa011374 protein of the invention

XX Sequence 364 AA;

Qy 352 LDMQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCSCRCRSLVELTPIAAVHG 411

Db 1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCSCRCRSLVELTPIAAVHG 60

Qy 412 SWGPRSPCSRSCGGVVTTRQCNPRPAPFGGRACVADLQAEVNTQACEKTQLEFMSQ 471

Db 61 SWGPRSPCSRSCGGVVTTRQCNPRPAPFGGRACVADLQAEVNTQACEKTQLEFMSQ 120

Qy 472 QCARTDQQLRSPGSGAFYHGAAPHSO-----GDAL-----CRMCR 531

Db 121 QCARTDQQLRSPGSGAFYHGAAPHSO-----GDAL-----CRMCR 180

Qy 532 MPSPGPRDGTLSLCVSGSCRTFGCDGRMDSQQVWDRQVCQVCGDSTCSFKRGSFTAGRAR 591

Db 181 MPSPGPRDGTLSLCVSGSCRTFGCDGRMDSQQVWDRQVCQVCGDSTCSFKRGSFTAGRAR 240

Qy 592 EYVTFLTVTNLTSVVIANHRPLFTHLAVRIGRYVVGKMSISPNTTYPSSLLEDGRVEY 651

Db 241 EYVTFLTVTNLTSVVIANHRPLFTHLAVRIGRYVVGKMSISPNTTYPSSLLEDGRVEY 300

Qy 652 RVALTEDRLPRLEIRIWGLQEDADIQV 680

Db 301 RVALTEDRLPRLEIRIWGLQEDADIQV 329

RESULT 15

ABP43989

ID ABP43989 standard; protein; 364 AA.

XX AC ABP43989;

XX 26-FEB-2003 (first entry)

XX Procollagen I N-proteinase.

XX Neuroprotective; immunomodulator; cancer; chromosome 9p34; cytostatic;

XX anti-inflammatory; gene therapy; nutritional supplement; wound; burn;

XX ulcer; Alzheimer's disease; Huntington's disease;

XX amyotrophic lateral sclerosis; autoimmune disorder; inflammation;

XX vulnerable.

XX Bos taurus.

XX WO200231111-A2.

XX 18-APR-2002.
XX PF 11-OCT-2001; 2001WO-US027760.
XX PR 12-OCT-2000; 2000US-00687527.
XX (HYSE-) HYSEQ INC.
XX PA
XX P1 Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX P1 Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX DR WPI: 2002-426278/45.
XX DR N-PSDB; ABQ61233.
XX PT New polypeptides and their encoded proteins, useful as nutritional
PT sources or supplements, or in gene therapy, particularly for treating
PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
PT inflammation.
XX PS Claim 20; SEQ ID # 892; 357pp + Sequence Listing; English.
XX CC The invention relates to 446 newly isolated polynucleotide sequences. The
CC activity of polynucleotides of the invention may be described as,
CC vulnary, neuroprotective, immunomodulator, cytostatic and anti-
CC inflammatory. Compositions comprising nucleic acids of the invention are
CC useful for treating a mammalian subject, or as nutritional sources or
CC supplements. These are useful in gene therapy, particularly for treating
CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
CC inflammation. The nucleic acids and polypeptides are also useful in
CC diagnostic and research methods. The sequences given in records ABP43544-
CC ABP43989 represent polypeptides encoded by polynucleotides of the
CC invention. NOTE: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 364 AA;

Query Match 23.28; Score 1799; DB 5; Length 364;
Best Local Similarity 99.78; Pred.No. 9.3e-127;
Matches 328; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 352 LDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 411
Db :|||||
1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 50

QY 412 SWGPRSPCSRSCGGVYVTRRRQNNPRPAFCGRACVGDALQAEKMTQACEKTLQEFMSQ 471
Db :|||||
61 SWGPRSPCSRSCGGVYVTRRRQNNPRPAFCGRACVGDALQAEKMTQACEKTLQEFMSQ 120

QY 472 QCARTDQQLRSPGGASFYHWGAAVPHSQDGLCRHWCRAIGSFIMKRGDSFLDGTTC 531
Db :|||||
121 QCARTDQQLRSPGGASFYHWGAAVPHSQDGLCRHWCRAIGSFIMKRGDSFLDGTTC 180

QY 532 MPSPGREDGTLSLCVSGSCRTFGCDGRMDSQVWDRQVCGDNDSTCSPKGSFTAGRAR 591
Db :|||||
181 MPSPGREDGTLSLCVSGSCRTFGCDGRMDSQVWDRQVCGDNDSTCSPKGSFTAGRAR 240

QY 592 EYVTFITVFNLSVYIANHRPLFTHLAVRIGGRVYVAGKMSISPNTTYPSSLEDGRVEY 651
Db :|||||
241 EYVTFITVFNLSVYIANHRPLFTHLAVRIGGRVYVAGKMSISPNTTYPSSLEDGRVEY 300

QY 652 RVALTEDRLPRLEIRIWGPLEQEDADIQV 680
Db :|||||
301 RVALTEDRLPRLEIRIWGPLEQEDADIQV 329

Search completed: March 13, 2004, 07:39:00
Job time : 216.238 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 112.1 Seconds
(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2

Perfect score: 7751

Sequence: 1 PGRPTRKAPSHSAPLGLA.....LQSWVPMQDPQSWKRGKT 1416

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/prodata/2/pubaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7751	100.0	1416	9	US-09-836-712-2
2	7604	98.1	1427	14	US-10-222-334-2
3	7596	98.0	1427	14	US-10-057-487-8
4	3609.5	46.6	842	14	US-10-222-334-4
5	3553.5	45.8	1120	9	US-09-781-080B-11
6	1799	23.2	365	9	US-09-978-979-5
7	1799	23.2	365	14	US-10-057-487-5
8	1459.5	18.8	738	9	US-09-978-979-6
9	1459.5	18.8	738	14	US-10-057-487-6
10	1261	16.3	242	9	US-09-978-979-1
11	1261	16.3	242	14	US-10-057-487-1
12	1190.5	15.4	1221	14	US-10-240-545A-2
13	1176.5	15.2	1629	9	US-09-972-467-2
14	1172.5	15.1	1916	15	US-10-274-639-10
15	1157	14.9	1882	9	US-09-918-171A-13

16	1148.5	14.8	1224	13	US-10-217-774-4	Sequence 4, Appli
17	1148.5	14.8	1224	14	US-10-296-616-2	Sequence 2, Appli
18	1141.5	14.7	1686	15	US-10-386-414-2	Sequence 2, Appli
19	1134.5	14.6	1690	9	US-09-788-043C-5	Sequence 5, Appli
20	1126	14.5	1593	11	US-09-981-151A-30	Sequence 30, Appli
21	1126	14.5	1593	15	US-10-295-027-1317	Sequence 1317, Ap
22	1125.5	14.5	2150	9	US-09-321-987B-2	Sequence 2, Appli
23	1118	14.4	2165	9	US-09-800-729-155	Sequence 155, App
24	1112.5	14.4	1907	9	US-09-938-330-25	Sequence 25, Appli
25	1072	13.8	1162	15	US-10-161-493-124	Sequence 124, App
26	1055	13.6	1123	15	US-10-120-801-45	Sequence 45, Appli
27	1051	13.6	203	9	US-09-781-080B-2	Sequence 2, Appli
28	1041.5	13.4	1057	14	US-10-188-869-10	Sequence 10, Appli
29	1041.5	13.4	1122	14	US-10-188-869-13	Sequence 13, Appli
30	1040.5	13.4	1081	15	US-10-120-801-46	Sequence 46, Appli
31	1039.5	13.4	1104	10	US-09-842-469-2	Sequence 2, Appli
32	1032.5	13.3	1104	10	US-09-842-469-4	Sequence 4, Appli
33	1036	13.2	1077	11	US-09-981-151A-31	Sequence 31, Appli
34	1026	13.2	1077	14	US-10-087-887-84	Sequence 84, Appli
35	1026	13.2	1077	15	US-10-120-801-49	Sequence 49, Appli
36	1024.5	13.2	1044	9	US-09-858-081-9	Sequence 9, Appli
37	1020	13.2	1145	14	US-10-188-869-20	Sequence 20, Appli
38	1017	13.1	1068	9	US-09-858-081-2	Sequence 2, Appli
39	1017	13.1	1133	9	US-09-858-068-2	Sequence 2, Appli
40	1004	13.0	854	11	US-09-981-151A-33	Sequence 33, Appli
41	957.5	12.4	1156	13	US-10-014-070-5	Sequence 5, Appli
42	957.5	12.4	1223	13	US-10-014-070-2	Sequence 2, Appli
43	957.5	12.4	1223	14	US-10-205-368-4	Sequence 4, Appli
44	957.5	12.4	1233	14	US-10-205-368-2	Sequence 2, Appli
45	957.5	12.4	1360	14	US-10-205-368-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-836-712-2

; Sequence 2, Application US/09836712

; Patent No. US20010049106A1

; GENERAL INFORMATION:

; APPLICANT: PFIZER INC.

; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND

; FILE REFERENCE: PC10851A

; CURRENT APPLICATION NUMBER: US/09/836,712

; CURRENT FILING DATE: 2001-04-17

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 2

; LENGTH: 1416

; TYPE: PRT

; ORGANISM: Human

US-09-836-712-2

Query Match 100.0%; Score 7751; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1416; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	PGRPTRKAPSHSAPLGLA	13	US-10-217-774-4	Sequence 4, Appli
Db	1	PGRPTRKAPSHSAPLGLA	13	US-10-296-616-2	Sequence 2, Appli
Qy	61	LQALEPOAVSSYLSPGAPL	61	US-09-781-080B-11	Sequence 11, Appli
Db	61	LQALEPOAVSSYLSPGAPL	61	US-10-057-487-5	Sequence 5, Appli
Qy	121	EDTRYVLTNLNIGAEILLR	121	US-10-057-487-6	Sequence 6, Appli
Db	121	EDTRYVLTNLNIGAEILLR	121	US-09-978-979-1	Sequence 1, Appli
Qy	181	SQTINPEDDTPGHADLVLY	181	US-10-274-639-10	Sequence 10, Appli
Db	181	SQTINPEDDTPGHADLVLY	181	US-09-918-171A-13	Sequence 13, Appli

241 LGVTTAHEIGHGFELEHGDAGSGCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAG 300
241 LGVTTAHEIGHGFELEHGDAGSGCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAG 300
301 RARCWMDPRPQPGSGAGHPDAPGLYYSANEORVAFGKAVACTFARHLDWCQALSC 360
301 RARCWMDPRPQPGSGAGHPDAPGLYYSANEORVAFGKAVACTFARHLDWCQALSC 360
361 HTDPLDQSSCSRLVPLLDGTEGVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCS 420
361 HTDPLDQSSCSRLVPLLDGTEGVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCS 420
421 RSCGGVWTRRQCNPNRPAFGRA CVGADLQAE MNTQACEKTQLEFMSQCCARTDGP 480
421 RSCGGVWTRRQCNPNRPAFGRA CVGADLQAE MNTQACEKTQLEFMSQCCARTDGP 480
481 LBSSFGGASFYHWGAAPHSQGDALCRHMCRAIGESFIMKRGDSFLDGTTCMPSPGREDG 540
481 LBSSFGGASFYHWGAAPHSQGDALCRHMCRAIGESFIMKRGDSFLDGTTCMPSPGREDG 540
541 TSLSCVSGSCRTFGCDGRWDSQVWDRQVCGGDNSTCSRKGSFTAGRAREYVTFVT 600
541 TSLSCVSGSCRTFGCDGRWDSQVWDRQVCGGDNSTCSRKGSFTAGRAREYVTFVT 600
601 PNLTSVYIANHRPLFTHLAVRIGRYVYVAGKMSISPNNTYPSLLEDGRVEYRVALTEDRL 660
601 PNLTSVYIANHRPLFTHLAVRIGRYVYVAGKMSISPNNTYPSLLEDGRVEYRVALTEDRL 660
661 PRLEIRIWGPLEOEDADIQVYRYGEBYGNLTRPDITFTYFQPKPRQAWWAAVGPSCV 720
661 PRLEIRIWGPLEOEDADIQVYRYGEBYGNLTRPDITFTYFQPKPRQAWWAAVGPSCV 720
721 SCGAGLRWNYSCLDQARKELVETVQCGSQOQPPAWPEACVLEPCPPVWAGDFGPCSAS 780
721 SCGAGLRWNYSCLDQARKELVETVQCGSQOQPPAWPEACVLEPCPPVWAGDFGPCSAS 780
781 CGGGLRERPVRCVEAOGSLKLTLPAPRCRAGAQQPAVALETNCNPPCPARVEVSEPSCT 840
781 CGGGLRERPVRCVEAOGSLKLTLPAPRCRAGAQQPAVALETNCNPPCPARVEVSEPSCT 840
841 SAGGAGLALENETCVPGADGLEAPVTEPGSVDEKLPAPEPCVGVSCPPGSHGLDATSAG 900
841 SAGGAGLALENETCVPGADGLEAPVTEPGSVDEKLPAPEPCVGVSCPPGSHGLDATSAG 900
901 EKAPSPWGSIRTAQAAHVTWPAAGSCSVSCGRGLMELRFLCMSALRVPVQBEELCGLAS 960
901 EKAPSPWGSIRTAQAAHVTWPAAGSCSVSCGRGLMELRFLCMSALRVPVQBEELCGLAS 960
961 KPGSRREVCOAPCPARWQYKLAACSVSCGRGVVRRILYCARHGEDDGEETLLDTQCG 1020
961 KPGSRREVCOAPCPARWQYKLAACSVSCGRGVVRRILYCARHGEDDGEETLLDTQCG 1020
1021 LRPPEQEAACSLBPCPPRWKMSLGPSCASCGLGTARRSVACVQLDQGDVEVDEBAACAA 1080
1021 LRPPEQEAACSLBPCPPRWKMSLGPSCASCGLGTARRSVACVQLDQGDVEVDEBAACAA 1080
1081 LVRPEASVPCLIADCTVRHVHTWMECSVSCGDIQRRRDTCLGPQAAPVADFCQHL 1140
1081 LVRPEASVPCLIADCTVRHVHTWMECSVSCGDIQRRRDTCLGPQAAPVADFCQHL 1140
1141 KPYTVRGWAGPCVQGTFSPLVPHHEAARPTTATPAGACRQHLEPTGTITDMGPGQA 1200
1141 KPYTVRGWAGPCVQGTFSPLVPHHEAARPTTATPAGACRQHLEPTGTITDMGPGQA 1200
1201 DCAVAIGRPLGEVVTLLVLESNLCSAGDMLLLWGLTWKRCRKLDDMTFSSKNTLVV 1260
1201 DCAVAIGRPLGEVVTLLVLESNLCSAGDMLLLWGLTWKRCRKLDDMTFSSKNTLVV 1260
1261 RQRCRPGGGVLLRYGSQLAPETTFYRECQMLFGPWGEIVSPSLSPATSNAGGCRFLINV 1320
1261 RQRCRPGGGVLLRYGSQLAPETTFYRECQMLFGPWGEIVSPSLSPATSNAGGCRFLINV 1320

1321 APHARIAHALATNMGAGTEGANASYILIRDTLSLRTTAFHGOQVLYWESSQAEMEPS 1380
1321 APHARIAHALATNMGAGTEGANASYILIRDTLSLRTTAFHGOQVLYWESSQAEMEPS 1380
1381 EGFLXAQASLRQCYWTLQSWVPQPMQPSQWKKEGT 1416
1381 EGFLXAQASLRQCYWTLQSWVPQPMQPSQWKKEGT 1416
RESULT 2
US-10-222-334-2
; Sequence 2, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
; FILE REFERENCE: UM-07288
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2
Query Match 98.1%; Score 7604; DB 14; Length 1427;
Best Local Similarity 97.6%; Pred. No. 0;
Matches 1393; Conservative 0; Mismatches 0; Indels 34; Gaps 1;
24 MEQRHPRARCPPLCVAGILACGFLGCGWSPSHFOQSCIQALEPOAVSSYLSGAPLKGRRP 83
1 MEQRHPRARCPPLCVAGILACGFLGCGWSPSHFOQSCIQALEPOAVSSYLSGAPLKGRRP 60
84 PSFGFORQORORRAAGGILHLELLVAVGPDVFOAQHQBDETERYVLTNLNIGALLRDPSL 143
61 PSFGFORQORORRAAGGILHLELLVAVGPDVFOAQHQBDETERYVLTNLNIGALLRDPSL 120
144 GAQFRVHLVKWVILTEPEGAPNTANLTSSLSVCGWSQTNPEDDTDPGHADLVLYITR 203
121 GAQFRVHLVKWVILTEPEGAPNTANLTSSLSVCGWSQTNPEDDTDPGHADLVLYITR 180
204 FDLELPDGNRQVRGVTLQGCASPTWSCLITETGFDLGVTHIAHIGHSFGLEHGDGAPGS 263
181 FDLELPDGNRQVRGVTLQGCASPTWSCLITETGFDLGVTHIAHIGHSFGLEHGDGAPGS 240
264 GCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAGRARCVWDPRPQPGSGAGHPDDAQ 323
241 GCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAGRARCVWDPRPQPGSGAGHPDDAQ 300
324 PGLYYSANEORVAFGKAVACTFARHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 383
301 PGLYYSANEORVAFGKAVACTFARHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 360
384 GVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCSRSGGVVTRRRQCNPRPAFG 443
361 GVEKWCCKGRCSLVELTPIAAVHGRWSSWGPRSPCSRSGGVVTRRRQCNPRPAFG 420
444 RACVAGDLQAE MNTQACEKTQLEFMSQCCARTDGP LRS SPGASFYHWGAAPHSQGD 503
421 RACVAGDLQAE MNTQACEKTQLEFMSQCCARTDGP LRS SPGASFYHWGAAPHSQGD 480
504 ALCRHMCRAIGESFIMKRGDSFLDGTTCMPSPGREDGTLSLCVSGSCRTFGCDGRWDSQ 563
481 ALCRHMCRAIGESFIMKRGDSFLDGTTCMPSPGREDGTLSLCVSGSCRTFGCDGRWDSQ 540
564 VWDRQVCGGDNSTCSRKGSFTAGRAREYVTFVTPLNLTSVYIANHRPLFTHLAVRIG 623

Db 541 VDRQVCGDNDSTSPKSGFTAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVRIG 600
Qy 624 GRVYVAGKMSISPTNTYPSLLEDGRVEYRVALTEDRLPRLEETIRWGPLOEDADIQVYRR 683
Db 601 GRVYVAGKMSISPTNTYPSLLEDGRVEYRVALTEDRLPRLEETIRWGPLOEDADIQVYRR 660
Qy 684 YGEYGNLTRPDITFTYFQPKPQAWAAVGRPCSVSCGAGLWVNYSCLDQARKELVE 743
Db 661 YGEYGNLTRPDITFTYFQPKPQAWAAVGRPCSVSCGAGLWVNYSCLDQARKELVE 720
Qy 744 TVOCQSQQPAPWPEACVLEPCPPYVAVGDFGPCSASCGGLRPRVRCVBAQSLKTL 803
Db 721 TVOCQSQQPAPWPEACVLEPCPPYVAVGDFGPCSASCGGLRPRVRCVBAQSLKTL 780
Qy 804 PPARCRAGAQQPAVALETNCPQPCPARWEVSEPSCTSSAGGAGLALENETCPVGADGLEA 863
Db 781 PPARCRAGAQQPAVALETNCPQPCPARWEVSEPSCTSSAGGAGLALENETCPVGADGLEA 840
Qy 864 PVTEGPGSVDEKLPAPETCPVCMGSCPPGKHLDATSAKEKAPSPWGSIRTGAAHVWTPA 923
Db 841 PVTEGPGSVDEKLPAPETCPVCMGSCPPGKHLDATSAKEKAPSPWGSIRTGAAHVWTPA 900
Qy 924 AGSCSVSCGRGLMELRFLCWDLSALRVVQBELCGLASKPGSRREVCOAVPCPARWQYKLA 983
Db 901 AGSCSVSCGRGLMELRFLCWDLSALRVVQBELCGLASKPGSRREVCOAVPCPARWQYKLA 960
Qy 984 ACSVSCGRGVVRRILYCARAHGEDDGEIILLDTCCQGLPRPEPOEACSLPCPPRWKVM 1043
Db 961 ACSVSCGRGVVRRILYCARAHGEDDGEIILLDTCCQGLPRPEPOEACSLPCPPRWKVM 1020
Qy 1044 LGPCASCSGLGTARRSVACVQLDQGVDEVEAACAALVRPEASVPCLIADCTYRWVHGT 1103
Db 1021 LGPCASCSGLGTARRSVACVQLDQGVDEVEAACAALVRPEASVPCLIADCTYRWVHGT 1080
Qy 1104 WMECSVSCGDGIQRRRTCLGPQQAQVPADFCOHLKPVTVRGCVWAGPCVQGGTPSLVP 1163
Db 1081 WMECSVSCGDGIQRRRTCLGPQQAQVPADFCOHLKPVTVRGCVWAGPCVQGGTPSLVP 1140
Qy 1164 HEEAAPGRTATPAG-----ACGROHLBPT 1189
Db 1141 HEEAAPGRTATPAGASLEWSQARGLLFSPAQPRLLPQENSVSSACGROHLBPT 1200
Qy 1190 GTIDMEGPGQADCAVAIGRPLGEVVTLRVLESSLNCSSAGDMLLLWGRITWRKMKCKLLDM 1249
Db 1201 GTIDMEGPGQADCAVAIGRPLGEVVTLRVLESSLNCSSAGDMLLLWGRITWRKMKCKLLDM 1260
Qy 1250 TFSKNTLTVRQRCRPGGGLLYRYSQAPETFYRECDMQLFGPWGEIVSPSLSPATS 1309
Db 1261 TFSKNTLTVRQRCRPGGGLLYRYSQAPETFYRECDMQLFGPWGEIVSPSLSPATS 1320
Qy 1310 NAGGCRLFINVAPHARIAIHALATNMAGTGANASYLIIIRDTHSLRTAFHQOQVLYWE 1369
Db 1321 NAGGCRLFINVAPHARIAIHALATNMAGTGANASYLIIIRDTHSLRTAFHQOQVLYWE 1380
Qy 1370 SESSQAEEMEFSEGLKQAQSLRGQYWTQLQSVPEMQDPQSMKKEGT 1416
Db 1381 SESSQAEEMEFSEGLKQAQSLRGQYWTQLQSVPEMQDPQSMKKEGT 1427

RESULT 3
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggreacanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1

; SEQ ID NO 8
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-057-487-8
Query Match 98.0%; Score 7596; DB 14; Length 1427;
Best Local Similarity 97.5%; Pred No. 0;
Matches 1392; Conservative 0; Mismatches 1; Indels 34; Gaps 1;
Qy 24 MHQRHPRARCPPLCVAGIILACGFLLCGWSHFQSCLOALEPOAVSSYLSFGAPLKGRP 83
Db 1 MHQRHPRARCPPLCVAGIILACGFLLCGWSHFQSCLOALEPOAVSSYLSFGAPLKGRP 60
Qy 84 PSFGFORQORQRAAGGILHLELLVAVGDPVFOAHQEDTERVYLTNINIGALLRDPDL 143
Db 61 PSFGFORQORQRAAGGILHLELLVAVGDPVFOAHQEDTERVYLTNINIGALLRDPDL 120
Qy 144 GAOFRVHLVKWVLTTPREGAPNITANLTSSILSVCGWSQTNIPEDDTPGHADLVYIIR 203
Db 121 GAOFRVHLVKWVLTTPREGAPNITANLTSSILSVCGWSQTNIPEDDTPGHADLVYIIR 180
Qy 204 FDELEPDGNGRQVGVTLQGGACSTWCLITETGFDLGVTTIAHEIGHSGFLEHGDGAPGS 263
Db 181 FDELEPDGNGRQVGVTLQGGACSTWCLITETGFDLGVTTIAHEIGHSGFLEHGDGAPGS 240
Qy 264 GCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAGRARCVWDPQPPQPSAGHPPDAQ 323
Db 241 GCGPSGHVWASDGAAPRAGLAWSPCSRQLLSLSAGRARCVWDPQPPQPSAGHPPDAQ 300
Qy 324 PGLYISANEQCRVAFGPKAVACTFAREHDMCQALSCHTDPLDQSSCSRLVPLLDGTSC 383
Db 301 PGLYISANEQCRVAFGPKAVACTFAREHDMCQALSCHTDPLDQSSCSRLVPLLDGTSC 360
Qy 384 GVEKWSKGRCSISVELTPTAAVHGRWSSGPRSPCSRSCGGVVTTRRCNNPRAFGG 443
Db 361 GVEKWSKGRCSISVELTPTAAVHGRWSSGPRSPCSRSCGGVVTTRRCNNPRAFGG 420
Qy 444 RACVGADLOAEMCNTQACEKTQLEFMSQCCARTDQQLRSPSPGGASFYHWGAAPVHSGD 503
Db 421 RACVGADLOAEMCNTQACEKTQLEFMSQCCARTDQQLRSPSPGGASFYHWGAAPVHSGD 480
Qy 504 ALCRHMCRAIGESFIMKRGDSFLDGTTRCMPSPGREDGTLSLCVSGSCRTFGDCGRMDSQ 563
Db 481 ALCRHMCRAIGESFIMKRGDSFLDGTTRCMPSPGREDGTLSLCVSGSCRTFGDCGRMDSQ 540
Qy 564 VDRQVCGDNDSTSPKSGFTAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVRIG 623
Db 541 VDRQVCGDNDSTSPKSGFTAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVRIG 600
Qy 624 GRVYVAGKMSISPTNTYPSLLEDGRVEYRVALTEDRLPRLEETIRWGPLOEDADIQVYRR 683
Db 601 GRVYVAGKMSISPTNTYPSLLEDGRVEYRVALTEDRLPRLEETIRWGPLOEDADIQVYRR 660
Qy 684 YGEYGNLTRPDITFTYFQPKPQAWAAVGRPCSVSCGAGLWVNYSCLDQARKELVE 743
Db 661 YGEYGNLTRPDITFTYFQPKPQAWAAVGRPCSVSCGAGLWVNYSCLDQARKELVE 720
Qy 744 TVOCQSQQPAPWPEACVLEPCPPYVAVGDFGPCSASCGGLRPRVRCVBAQSLKTL 803
Db 721 TVOCQSQQPAPWPEACVLEPCPPYVAVGDFGPCSASCGGLRPRVRCVBAQSLKTL 780
Qy 804 PPARCRAGAQQPAVALETNCPQPCPARWEVSEPSCTSSAGGAGLALENETCPVGADGLEA 863
Db 781 PPARCRAGAQQPAVALETNCPQPCPARWEVSEPSCTSSAGGAGLALENETCPVGADGLEA 840
Qy 864 PVTEGPGSVDEKLPAPETCPVCMGSCPPGKHLDATSAKEKAPSPWGSIRTGAAHVWTPA 923
Db 841 PVTEGPGSVDEKLPAPETCPVCMGSCPPGKHLDATSAKEKAPSPWGSIRTGAAHVWTPA 900
Qy 924 AGSCSVSCGRGLMELRFLCWDLSALRVVQBELCGLASKPGSRREVCOAVPCPARWQYKLA 983
Db 901 AGSCSVSCGRGLMELRFLCWDLSALRVVQBELCGLASKPGSRREVCOAVPCPARWQYKLA 960

```
QY 984 ACSVSCGRGVRIRILYCARAHGDDGEBEILLDTQCQGLPRPEPOEACSLFPPRPWKVMS 1043
DB 961 ACSVSCGRGVRIRILYCARAHGDDGEBEILLDTQCQGLPRPEPOEACSLFPPRPWKVMS 1020
QY 1044 LGPCASCGLTARRSVACVQLDQGGQVDEDAACAAALVRPEASVPCLIADCTYRWHVGT 1103
DB 1021 LGPCASCGLTARRSVACVQLDQGGQVDEDAACAAALVRPEASVPCLIADCTYRWHVGT 1080
QY 1104 WNECSVSCGDIORRDTCLGPQAQAVPADFCOHLKPKVTVRCWAGPCVGGTSPSLVP 1163
DB 1081 WNECSVSCGDIORRDTCLGPQAQAVPADFCOHLKPKVTVRCWAGPCVGGTSPSLVP 1140
QY 1164 HEAAAPGRTTAPAG-----ACGRQHLEPT 1189
DB 1141 HEAAAPGRTTAPAGASLWSQARGLLPSPAPQPRILLPQPOENSVOSSACGRQHLEPT 1200
QY 1190 GTIDMRGPGQADCAVAIGRPLGEVVTILRVLESSLNCAGDMLLWGLRTWRKCRKLLDM 1249
DB 1201 GTIDMRGPGQADCAVAIGRPLGEVVTILRVLESSLNCAGDMLLWGLRTWRKCRKLLDM 1260
QY 1250 TFSKNTNLTVVRQCRPGGVLRYGSQLAPETFYRECDMLQFGPWGEIYVSPSLSPATS 1309
DB 1261 TFSKNTNLTVVRQCRPGGVLRYGSQLAPETFYRECDMLQFGPWGEIYVSPSLSPATS 1320
QY 1310 NAGGCRFLFNVAHARITAHALATNMAGTEGANASYILIRDTSLRTTAFHGGQVLYWE 1369
DB 1321 NAGGCRFLFNVAHARITAHALATNMAGTEGANASYILIRDTSLRTTAFHGGQVLYWE 1380
QY 1370 SESSQAESEFSEGLKAQASLRGQVWTLQSVPMQDPQSWKKEGT 1416
DB 1381 SESSQAESEFSEGLKAQASLRGQVWTLQSVPMQDPQSWKKEGT 1427

RESULT 4
US-10-222-334-4
; Sequence 4, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Teai, Han-Mou
; TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof.
; FILE REFERENCE: US-07288
; CURRENT APPLICATION NUMBER: US/10/222,334
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-4

Query Match 46.6%; Score 3609.5; DB 14; Length 842;
Best Local Similarity 77.1%; Pred. No. 2.5e-257;
Matches 712; Conservative 20; Mismatches 90; Indels 101; Gaps 16;

QY 24 MHQHPRARCPLCVAGILACGLFLGCGWPSHFQSCLOALEPQAVSSYLSFGAPLKGRP 83
DB 1 MHQHPRARCPLCVAGILACGLFLGCGWPSHFQSCLOALEPQAVSSYLSFGAPLKGRP 60
QY 84 PSPGFQQRORRAAGILHLELVAVGPDVFOAQHEDTERVLTNLTNLTGAEILLRDPDL 143
DB 61 PSPGFQQRORRAAGILHLELVAVGPDVFOAQHEDTERVLTNLTNLTGAEILLRDPDL 120
QY 144 GAQPRVHLVKVILTEPGAPNITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITR 203
DB 121 GAQPRVHLVKVILTEPGAPNITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITR 190
QY 204 FDELPDGNRQVRGVTQLGGACSTWCLITDTPGLGVITIAHGHSGFLEHGDGAPGS 263
```

```
181 FDELPDGNRQVRGVTQLGGACSTWCLITDTPGLGVITIAHGHSGFLEHGDGAPGS 240
264 GCGPSGHVNASDGAAPRAGLAWSPCSRRLQLLSLSAGRARCVMPPRPPQPSAGHPDDAQ 323
241 GCGPSGHVNASDGAAPRAGLAWSPCSRRLQLLSLSAGRARCVMPPRPPQPSAGHPDDAQ 300
324 PGLYYSANEQCRVAFQPKAVACTFAREHLMDCQALSCHTDPLDQSSCSRLLVPLLDGTREC 383
301 PGLYYSANEQCRVAFQPKAVACTFAREHLMDCQALSCHTDPLDQSSCSRLLVPLLDGTREC 360
384 GVEKWCCKRCRSLVLTPIAAVHGRWSSWSPRSCPSRSCGGVVTRRQCNPNRPAFGG 443
361 GVEKWCCKRCRSLVLTPIAAVHGRWSSWSPRSCPSRSCGGVVTRRQCNPNRPAFGG 420
444 RACVGLADLAEMCNTQACEKTQLEFMSQOQCARTDGQRLSRSPGGASFYHWAAGVPHSQD 503
421 RACVGLADLAEMCNTQACEKTQLEFMSQOQCARTDGQRLSRSPGGASFYHWAAGVPHSQD 480
504 ALCRHMCRAIGESFIMKRGDSFLDGTRCMPSPGREDGTLSLCVSGSCRTFGCDGRMDSQ 563
481 ALCRHMCRAIGESFIMKRGDSFLDGTRCMPSPGREDGTLSLCVSGSCRTFGCDGRMDSQ 540
564 VMDRCQVCGDNDSTCSPRKGSTAGRAREYVTFVTTPNLTSVYIANHRPLFTHLAVRIG 623
541 VMDRCQVCGDNDSTCSPRKGSTAGRAREYVTFVTTPNLTSVYIANHRPLFTHLAVRIG 600
624 GRVYVAGKMSISNTTYPSSLEDDGRVEYRVALTEDRLPRLEIRIWGLPQEDADIQV-- 681
601 GRVYVAGKMSISNTTYPSSLEDDGRVEYRVALTEDRLPRLEIRIWGLPQEDADIQVFC 660
682 RRYGEEVGNL-----TRPDITFTYFQPKPQAAVMAAVRGPSCVSCGAGLEWVNYSCLD 735
661 RFTGGWARSNATSPATSPSTSLSHGRP-----GCGP-----LCVG 698
736 QARKELVETVQCGSQPPPAWPEACVLEPCPPYVAVGDFGSCSACGGGLRERPRVCVEA 795
699 PARAAAL-----GKLQLPQGPQGVGDCPV-----PREPAATSVARGLRARTLPSSLG 746
796 QGSLLKTLPPARCAGAAQPAVALETCPQPCPARWEVSEPSSTCSAGGAGLALENETCV 855
747 GGBL-----RPMQRLWG-----W-----PAGAASALRG----- 770
856 PGADGLEAPVTEGPGSVDEKL-PAPEPCVGVSCPPGWHGLDATSAGEKAPSPWGSIRTGA 914
771 -----GPGQPPEDIAPSVQSRGPAASCAGNLQPPA-----LPCQVGGVR--A 812
915 QAAHV--WT-PAAGSCSVSCGRG 934
813 QLMHISWWSRPLGERDL-CARG 834

RESULT 5
US-09-781-080B-11
; Sequence 11, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: YAMAMOTO, GAYLE
; TITLE OF INVENTION: Anti-Angiogenic Intestinal Peptides,
; FILE REFERENCE: 99-82
; CURRENT APPLICATION NUMBER: US/09/781,080B
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 1120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
```

LOCATION: (1)....(1120)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-781-0808-11

Query Match 45.8%; Score 3553.5; DB 9; Length 1120;
Best Local Similarity 54.2%; Pred. No. 5e-253;
Matches 766; Conservative 21; Mismatches 108; Indels 517; Gaps 25;

QY 8 KAPSHAPLILGLALRMHQRHRCRPPCLVAGILACGFLGCGWSPHFQSCIQALEPQ 67
DB 15 KAPSHAPLILGLALRMHQRHRCRPPCLVAGILACGFLGCGWSPHFQSCIQALEPQ 74
QY 68 AVSSYLSGAPLGRPPSPGFQQRQRRRAAGGILHLLVAVGPDVFQAHQEDTERV 127
DB 75 AVSSYLSGAPLGRPPSPGFQQRQRRRAAGGILHLLVAVGPDVFQAHQEDTERV 134
QY 128 LTNLNTGAEILLRDPISLGAQRVHLVKVILTEPEGAPNITANLTSSLSVCGWSQTINPE 187
DB 135 LTNLNTGAEILLRDPISLGAQRVHLVKVILTEPEGAPNITANLTSSLSVCGWSQTINPE 194
QY 188 DTDDPGHADLVLYITRFDLELPGNRQVRGVYQLGGACSPWMSCLITDGTGDLGVITIAH 247
DB 195 DTDDPGHADLVLYITRFDLELPGNRQVRGVYQLGGACSPWMSCLITDGTGDLGVITIAH 254
QY 248 EIGHSPGLEHGDGAPGSCGSHVWASDGAAPRAGLAWSPCSRROLLS-----LISAGRAR 303
DB 255 EIGHSPGLEHGDGAPGSCGSHVWASDGAAPRAGLAWSPCSRROLLSAGPQALRVGTRR 314
QY 304 CYNDDPRPQGSAGHPDAPQ-GLYSGANEOCRVAFGPKAVACTFAREHLDN---COALS 359
DB 315 GL-----KPGFTGAPAGWSLGLYISANEQCHVAFPGG-----CRUHLRQGAPCOALS 363
QY 360 CHTDPLDQSSCSRLVPLLDGTECGVKWCSKGRCSRLVELTPIAAVHGRMSSWGPRSPC 419
DB 364 CHTDPLDQSSCSRLVPLLDGTECGVEK-----VHGRMSSWGPRSPC 405
QY 420 SRSCGGVTRRQCNN-----PRPAFGGRACVGDADLQBMCMNTQACEKTQLEBFMSQ 472
DB 406 SRSCGG-----CGHQAEEVQPPVQGGFLCQGMKLGWGTCPSCSVQDTLFHS-- 455
QY 473 CARTDQPLRSSFGGASFYHGAAPVPHSGDALCRHMCRAIGESFIMKRGD-SFLDGTGC 531
DB 456 -----ALP---GDALCRHMCRAIGESXHBAAWETASSNGTC 489
QY 532 MPSPGREDGTLSLCVSGSCRTFGCDGRMDSQVWDRQVQCGDNDSTCSPKGSFTAGRAR 591
DB 490 MPSPGREDGTLSLCVSGSCRVGCDGRMDSQVWDRQVQCGDNDSTCSPKGSFTAGRAR 549
QY 592 EYVTFITVNTLTSVYIANHRPLFTHLAVRIGRYVAVAGKMSISPTNTYPSLLEDGRVEY 651
DB 550 EYVTFITVNTLTSVYIANHRPLFTHLAVRIGRYVAVAGKMSISPTNTYPSLLEDGRVEY 609
QY 652 RVALTEDRLRLBEIRIWGPLEQEDADIQVYRRYGEEYGNLRTPOITFTYFQPKPQAWVM 711
DB 610 ----- 609
QY 712 AAVRGPCSVCGAGLRVWVNSCLDQARKELVETVQCGSQPPANPEACVLEPCPYWAV 771
DB 610 -----QCVKKQIPGS-----SAYS LN 625
QY 772 GDFGPCASCGGLRERPVRCVEAQGSLKTLPPARCAGAQPPAVALTECNPPQCPAEM 831
DB 626 QDFPVL-----GLNRV-----TWGLR-----HW 646
QY 832 EVSEPPSCTSAGGAGLALNETCVPGADLEAPVTEGFSVDEKLPAPPCVGMSCPPGW 891
DB 647 -----PDVGIE----- 652
QY 892 GHLDATGAGEKAPGSPWGSIRTAQAHAHVTWPAAGSCSVSCGRLMELRFLCMDSALRVPV 951
DB 653 -----GAGLMELRFLCMDSALRVPV 672
QY 952 QBEELCGLASKPGSRREVQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGE 1011

DB 673 QBEELCGLASKPGSRREVQAVPCPARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGE 732
QY 1012 ILDDTQCQGLRPEPEQACSLPCPPRWKVMNSLGFCSASCGLGTARRSVACVQLDQGDV 1071
DB 733 ILDDTQCQGLRPEPEQACSLPCPPRWKVMNSLGFCSASCGLGTARRSVACVQLDQGDV 792
QY 1072 EVDEAACALVRPEASVPLIADCTYRHHVGTWMECSVSCGDIQORRDTCLGPQAQAPV 1131
DB 793 EVDEAACALVAAR----- 806
QY 1132 PADFCQHLKPRVTVRGWAGPCVCGQTPSLVPHEEAAAPGRTTATPAGACGRQHLEPTGT 1191
DB 807 -----GQLSPVSLP-----TATAGM-----LAPGWR 828
QY 1192 IDMRGPOADCAVAIGRPLGEVWTLRVLESSLNC-AGDMLLLWGLTWKRCRKLDDMT 1250
DB 829 RGSAGP-----SLPCRFLGDMLLWGLTWKRCRKLDDMT 864
QY 1251 FSSKTTLVVRQRCRPGCGVLLRYGSQLADETFYRECDMQLFGPWGEIVSLSLSPATSN 1310
DB 865 FSSKTTLV----- 873
QY 1311 AGCRLFINVAPHARIAHALATNMGAGTEGANASYILIRDTHTSLRTTAFH-GQOVLYWE 1369
DB 874 -----IRDTHTSLRTTAFHRAAGALTGS 895
QY 1370 SESSQEMEPSEGFKAQASIRGOYWTIQSWV 1401
DB 896 QRAARLEMESEGFKAQASIRGOYWTIQSWL 927

RESULT 6
US-09-978-979-5
Sequence 5, Application US/09978979
Patent No. US20020151702A1
GENERAL INFORMATION:
APPLICANT: Racie, Lisa, A.
Twine, Natalie, C.
Agostino, Michael, J.
Wolfman, Neil
Morris, Elisabeth
TITLE OF INVENTION: Aggreacanase Molecules
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Home Products Corporation
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/978,979
FILING DATE: 16-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: GI 5435P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5;
US-09-978-979-5

Query Match 23.2%; Score 1799; DB 9; Length 365;
Best Local Similarity 99.7%; Pred. No. 2.8e-124;
Matches 328; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 352 LDMCQALSCHTDPLDQSSCRLLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 411
DB 1 :
QY 412 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 471
DB 61 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 120
QY 472 QCARTDQQLRSSPGGASFYHWGAAPVPHSGDALCRHMCRAIGESFIMKRGDSFLDGTGTC 531
DB 121 QCARTDQQLRSSPGGASFYHWGAAPVPHSGDALCRHMCRAIGESFIMKRGDSFLDGTGTC 180
QY 532 MPSPGPRDGTLSLCVSGSCRTFCGCDGRMDSQQVWDRQVCGDNDSTCSPRKGSTAGRAR 591
DB 181 MPSPGPRDGTLSLCVSGSCRTFCGCDGRMDSQQVWDRQVCGDNDSTCSPRKGSTAGRAR 240
QY 592 EYVTLVTPNLTSVYIANHRPLFTHLAVRIGGRVYVWAGKMSISPNNTYPSLLEDGRVEY 651
DB 241 EYVTLVTPNLTSVYIANHRPLFTHLAVRIGGRVYVWAGKMSISPNNTYPSLLEDGRVEY 300
QY 652 RVALTEDRLPRLEIRINGWPLQEDADIQV 680
DB 301 RVALTEDRLPRLEIRINGWPLQEDADIQV 329

TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5;
US-09-978-979-5

Query Match 23.2%; Score 1799; DB 9; Length 365;
Best Local Similarity 99.7%; Pred. No. 2.8e-124;
Matches 328; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 352 LDMCQALSCHTDPLDQSSCRLLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 411
DB 1 :
QY 412 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 471
DB 61 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 120
QY 472 QCARTDQQLRSSPGGASFYHWGAAPVPHSGDALCRHMCRAIGESFIMKRGDSFLDGTGTC 531
DB 121 QCARTDQQLRSSPGGASFYHWGAAPVPHSGDALCRHMCRAIGESFIMKRGDSFLDGTGTC 180
QY 532 MPSPGPRDGTLSLCVSGSCRTFCGCDGRMDSQQVWDRQVCGDNDSTCSPRKGSTAGRAR 591
DB 181 MPSPGPRDGTLSLCVSGSCRTFCGCDGRMDSQQVWDRQVCGDNDSTCSPRKGSTAGRAR 240
QY 592 EYVTLVTPNLTSVYIANHRPLFTHLAVRIGGRVYVWAGKMSISPNNTYPSLLEDGRVEY 651
DB 241 EYVTLVTPNLTSVYIANHRPLFTHLAVRIGGRVYVWAGKMSISPNNTYPSLLEDGRVEY 300
QY 652 RVALTEDRLPRLEIRINGWPLQEDADIQV 680
DB 301 RVALTEDRLPRLEIRINGWPLQEDADIQV 329

RESULT 7
US-10-057-487-5
; Sequence 5, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrucanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 365
; TYPE: PRT
; ORGANISM: homo sapiens
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: unknown amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (365)..(365)
; OTHER INFORMATION: unknown amino acid
US-10-057-487-5

Query Match 23.2%; Score 1799; DB 14; Length 365;
Best Local Similarity 99.7%; Pred. No. 2.8e-124;
Matches 328; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 352 LDMCQALSCHTDPLDQSSCRLLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 411
DB 1 :
QY 412 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 471

RESULT 8
US-09-978-979-6
; Sequence 6, Application US/09978979
; Patent No. US20020151702A1
; GENERAL INFORMATION:
; APPLICANT: Racie, Lisa, A.
; Twine, Natalie, C.
; Agostino, Michael, J.
; Wolfman, Neil
; Morris, Elisabeth
; TITLE OF INVENTION: Aggrucanase Molecules
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/978,979
; FILING DATE: 16-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/60/241,469
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,345
; REFERENCE/DOCKET NUMBER: GI 5435p
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 660-5000
; TELEFAX: (973) 683-4117
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 738 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-978-979-6

Query Match 18.8%; Score 1459.5; DB 9; Length 738;
Best Local Similarity 39.4%; Pred. No. 7.9e-99;
Matches 374; Conservative 34; Mismatches 163; Indels 379; Gaps 32;

QY 252 SFGLEHDAFGSGCGPSCGHVWASDGAAP-----RAGLAWSPC 288

RESULT 7
US-10-057-487-5
; Sequence 5, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrucanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 365
; TYPE: PRT
; ORGANISM: homo sapiens
; NAME/KEY: MISC FEATURE
; OTHER INFORMATION: unknown amino acid
; NAME/KEY: MISC FEATURE
; LOCATION: (365)..(365)
; OTHER INFORMATION: unknown amino acid
US-10-057-487-5

Query Match 23.2%; Score 1799; DB 14; Length 365;
Best Local Similarity 99.7%; Pred. No. 2.8e-124;
Matches 328; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 352 LDMCQALSCHTDPLDQSSCRLLVPLLDGTECGVEKWCCKRCRSLVELTPIAAVHGWS 411
DB 1 :
QY 412 SWGPRSPCSRSCGGGVVTRRQCNPRPAFGGRACVADLAQEMCNTQACEKTQLEFMSQ 471

Db 1 SFGLHDGAPGSCGPGSHVMASERRRPPAPSPGPPAAAGSCXACSDPSLRRSLCWPPT 60
Qy 289 SR-----ROLLSLLSAGRA----- 302
Db 61 SAPAGALVLPKSRLLVGGAGRELLPLTKGHASKRPHFRAHSSVPPPGVHPGTEPGL 120
Qy 303 -----RCVWDPRPQPGSAGHPDPAQGLYYSANEQCRVAFGPKAVACTFAREHL 352
Db 121 SRALSQMTGALVWDPRPQPGSAGHPDPAQGLYYSANEQCRVAFGPKAVACTFAREH- 179
Qy 353 DMQALSCHTDPLDQSSCSRLVPLLDGTEGVEKWCCKGRCSRLVELTPIAAVHGRWSS 412
Db 180 -----LVSLPAVANDWLX 192
Qy 413 WGRPSRCSRSCGGVTVRRQCNPRPAFGGRACVADLQAEKMCNTQACEKTLQEFMSQ 472
Db 193 GPSASPSRRP-----PKRAMICARPSA-----TQTRWTKAAAAAASFSLSMGQN 237
Qy 473 CARTDQCP-LRSPGASGFVHGAAPHSDALCRHMCRAIGSFIMKRGDSFLDGTGRC 531
Db 238 VAWRSGAPRAAAP-----WNSXPEXQ-----C 261
Qy 532 MFSGPREDGTLSCVSGSCRTFGCDGRMDQVDRQVCGDNDSTCSPKSGFTAGRAR 591
Db 262 MGAG-----LAGVPEVL---APAPAEVW-----SPGGSAT----- 290
Qy 592 EYVFTLVTPLNLTSVYIANHRLPFLTHLAVRIGRYVAVAGKMSISBNTTTPSYLLEDGRVEY 651
Db 291 -----TPDLP-----LGGVHVL----- 302
Qy 652 RVALTEDRLPRLBEIR-IWGPLEQDADIQYRRYGEYGNLTPDITFTYFPKPRQAW 709
Db 303 -VLTSRPRCATLPAKPPSPSSCNSAPGPTASRAPP---LAAP-----PSTTGLV 350
Qy 710 VMAAVGCP-CSVCGAGLRWYNSCLDQARKVELTVCQSGSQPPAWPEACVLEPCPPY 768
Db 351 LVHTAKGMLCADTCAGPLA-----RASSXSVETASSMG-----PGVCQVAP----- 391
Qy 769 WAVDPRP-----CSASCG-----GGLRER---PVRV-----EAQSLLK----- 801
Db 392 ---GRTPGXACVCRAAAGHLAVWGVTPSRYGTGARCVCVGTARAAGRALSQLAERNM 448
Qy 802 ---TLPPA-----RCRAGQ-----OPA---VALETCN 823
Db 449 SRFXPQPTXVSTLPTTGLSSHTRWGSSEAMSLGRXASPLTPPTPPSRMVVSSTEW 508
Qy 824 POP---CPA-----RWEVSPSSCTSSAGGAGLALENETCPVPGADGLEAP 864
Db 509 PSRTGCPAWERSAGSDPSRKMLTSRWEVSPSSCTSSAGGAGLALENETCPVPGADGLEAP 568
Qy 865 VTEGPGSVDEKLPAPEPCVGMSCPPGWHLDATSGAGXAPSPWGSIRTCQAQAAHVWTPAA 924
Db 569 VTEGPGSVDEKLPAPEPCVGMSCPPGWHLDATSGAGXAPSPWGSIRTCQAQAAHVWTPAA 628
Qy 925 GSCSVSCGRGLMELRFLCMLSALRVPVQEEELCGLASKPGSRREVCOAVPCPARWQYKLA 984
Db 629 GSCSVSCGRGLMELRFLCMLSALRVPVQEEELCGLASKPGSRREVCOAVPCPARWQYKLA 688
Qy 985 CSVSCGGRVRRILYCARAHGEDDEBELLTTOCOGLPRPEPOEACSELP 1034
Db 689 CSVSCGGRVRRILYCARAHGEDDEBELLTTOCOGLPRPEPOEACSELP 738

RESULT 9

US-10-057-487-6
; Sequence 6, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrucanase Molecules
; FILE REFERENCES: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25

; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 738
; TYPE: PRT
; ORGANISM: homo sapien
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (43)..(43)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (192)..(192)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (255)..(255)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (358)..(358)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (374)..(374)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (397)..(397)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (452)..(452)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (458)..(458)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (475)..(475)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (487)..(487)
; OTHER INFORMATION: unknown amino acid
; US-10-057-487-6

Query Match

Best Local Similarity 18.8%; Score 1459.5; DB 14; Length 738;
Matches 374; Conservative 34; Mismatches 163; Indels 379; Gaps 32;

Qy 252 SFGLHDGAPGSCGPGSHVMASDGAAP-----RAGLAWSPC 288
Db 1 SFGLHDGAPGSCGPGSHVMASRRRPPAPSPGPPAAAGSCXACSDPSLRRSLCWPPT 60
Qy 289 SR-----ROLLSLLSAGRA----- 302
Db 61 SAPAGALVLPKSRLLVGGAGRELLPLTKGHASKRPHFRAHSSVPPPGVHPGTEPGL 120
Qy 303 -----RCVWDPRPQPGSAGHPDPAQGLYYSANEQCRVAFGPKAVACTFAREHL 352
Db 121 SRALSQMTGALVWDPRPQPGSAGHPDPAQGLYYSANEQCRVAFGPKAVACTFAREH- 179
Qy 353 DMQALSCHTDPLDQSSCSRLVPLLDGTEGVEKWCCKGRCSRLVELTPIAAVHGRWSS 412
Db 180 -----LVSLPAVANDWLX 192
Qy 413 WGRPSRCSRSCGGVTVRRQCNPRPAFGGRACVADLQAEKMCNTQACEKTLQEFMSQ 472
Db 193 GPSASPSRRP-----PKRAMICARPSA-----TQTRWTKAAAAAASFSLSMGQN 237

473 CARTGQP-LRSPGGSFTHGAAVPHSQDGLCRHMCRAIGSFIMKRGDSFLDGTGRC 531
238 VAMRSGAPRAAAP-----MWSXPPXQ-----C 261
532 MPSGREDDTSLCVSSGSCRTFCDCGDMDSQVWDRQVCGGDNSTCSPRKGSFTAGRAR 591
262 MGAG-----LAGVEVL--APAPAEVW-----SPGGGSAT----- 290
592 EYVTFITVTPNLTSVYIANRPLFTHLAVRIGGRYVAGKMSISPNTTYPSSLLEDGRVEY 651
291 -----TPDLP-----LGGVHL----- 302
652 RVALTEDRLPRLEIR--IWPIQEDADIQVYRYGEEYGNLTPDITFTFYQPKPROAM 709
303 -VLTSPRCATLRPARPSSWSSCKNSAPGPTASRCAPP--LAAP-----ESTIGVL 350
710 VNAAVRGP-CSVSCGAGLRWNWSCLDQARKELVETVQCQSQPPPAWPEACVLEPCPPY 768
351 LYHTAKGMLCADTCAGPLA-----RASSXSVEATSSMG-----PGVCQVAP----- 391
769 WAVGDFGP---CSASCG-----GGLRER---PVRGV-----EAQSLK----- 801
392 ---GRTPXACVCRAGHLLAVMVGWTPSRVGTGARCVCVGTGTTARAAGRALSQLAERNM 448
802 ---TLPPA-----RCRAGAO-----OPA---VALETCN 823
449 SRFXQLPPYXVSTLPTTGLSSHTWRGSEGMWLGEXASPLTPPTPSRWVVSSTEW 508
824 PQP---CPA-----RWEVSPSCTSAGGAGLALENETCVPGADGLEAP 864
509 PSPTGCPAWRRSASGDPKMLTSRWEVSPSCTSAGGAGLALENETCVPGADGLEAP 568
865 VTEGPGSVDEKLPAPEPCVGMSCPPGHLDTAGKAPSPWGSIRTGAAAHVWTPAA 924
569 VTEGPGSVDEKLPAPEPCVGMSCPPGHLDTAGKAPSPWGSIRTGAAAHVWTPAA 628
925 GSCSVCGRLMELRFLCMDSALRVPVQEEICGLASKPGSRREVCOAVPCPARWOYKLA 984
629 GSCSVCGRLMELRFLCMDSALRVPVQEEICGLASKPGSRREVCOAVPCPARWOYKLA 688
985 CSVSCGGRVVRILYCARAHGEDDGEIILLDTQCQGLPRPPOEACSLP 1034
689 CSVSCGGRVVRILYCARAHGEDDGEIILLDTQCQGLPRPPOEACSLP 738

RESULT 10

US-09-978-979-1

Sequence 1, Application US/09978979

Patent No. US20020151702A1

GENERAL INFORMATION:

APPLICANT: Racine, Lisa, A.

Twine, Natalie, C.

Agostino, Michael, J.

Wolfman, Neil

Morris, Elisabeth

TITLE OF INVENTION: Aggrecanase Molecules

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: American Home Products Corporation

STREET: One Campus Drive

CITY: Parsippany

STATE: New Jersey

COUNTRY: USA

ZIP: 07054

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/978,979

FILING DATE: 16-Oct-2001

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: CI 5435P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-978-979-1
Query Match 16.3%; Score 1261; DB 9; Length 242;
Best Local Similarity 100.0%; Pred. No. 7.6e-85;
Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 SCLOALEPQAVSSYLSFGAPLKGRPPSPGFQQRORORRAAGGILHLELLVAVGPDVFOA 118
Db 3 SCLOALEPQAVSSYLSFGAPLKGRPPSPGFQQRORORRAAGGILHLELLVAVGPDVFOA 62
QY 119 HQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKNVILTEPEGAPNITANTLSSLSVC 178
Db 63 HQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKNVILTEPEGAPNITANTLSSLSVC 122
QY 179 GMSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTLQGGACSTWASCLITDGTG 238
Db 123 GMSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTLQGGACSTWASCLITDGTG 182
QY 239 FDLGVITAIHIGHSFGLHEDGAPGCGGPGSHVMSDGAAPRAGLAWSPCSRRLLSLL 297
Db 183 FDLGVITAIHIGHSFGLHEDGAPGCGGPGSHVMSDGAAPRAGLAWSPCSRRLLSLL 241

RESULT 11

US-10-057-487-1

Sequence 1, Application US/10057487

Publication No. US20030105313A1

GENERAL INFORMATION:

APPLICANT: Wyeth

TITLE OF INVENTION: Aggrecanase Molecules

FILE REFERENCE: 08702.0073

CURRENT APPLICATION NUMBER: US/10/057,487

CURRENT FILING DATE: 2002-01-25

PRIOR APPLICATION NUMBER: 60/241,469

PRIOR FILING DATE: 2000-10-18

NUMBER OF SEQ ID NOS: 8

SOFTWARE: Patent in version 3.1

SEQ ID NO 1

LENGTH: 242

TYPE: PRT

ORGANISM: Homo sapiens

US-10-057-487-1

Query Match 16.3%; Score 1261; DB 14; Length 242;

Best Local Similarity 100.0%; Pred. No. 7.6e-85;

Matches 239; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 SCLOALEPQAVSSYLSFGAPLKGRPPSPGFQQRORORRAAGGILHLELLVAVGPDVFOA 118

Db 3 SCLOALEPQAVSSYLSFGAPLKGRPPSPGFQQRORORRAAGGILHLELLVAVGPDVFOA 62

QY 119 HQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKNVILTEPEGAPNITANTLSSLSVC 178

Db 63 HQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKNVILTEPEGAPNITANTLSSLSVC 122

QY 179 GMSQTINPEDDDTPGHADLVLYITRFDLELPGNRQVRGVTVQLGGACSPWTSCLITEDTG 238
Db 123 GMSQTINPEDDDTPGHADLVLYITRFDLELPGNRQVRGVTVQLGGACSPWTSCLITEDTG 182
QY 239 FDLGVTTIAHEIGHSGFLEHGDGPGSGCGPSGHVNASDGAAPRAGLAWSPCSRRLQLSLL 297
Db 183 FDLGVTTIAHEIGHSGFLEHGDGPGSGCGPSGHVNASDGAAPRAGLAWSPCSRRLQLSLL 241

RESULT 12
US-10-240-545A-2
; Sequence 2, Application US/10240545A
; Publication No. US20030185828A1
; GENERAL INFORMATION:
; APPLICANT: Yamanouchi Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: No. US20030185828A1e1 aggrecanase
; FILE REFERENCE: 08959.0002
; CURRENT APPLICATION NUMBER: US/10/240,545A
; PRIOR FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/JP01/11033
; PRIOR FILING DATE: 2001-12-17
; PRIOR APPLICATION NUMBER: JP 2000-384300
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 26
; SEQ ID NO 2
; LENGTH: 1221
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-240-545A-2

Query Match 15.4%; Score 1190.5; DB 14; Length 1221;
Best Local Similarity 28.4%; Pred No. 1.1e-78;
Matches 312; Conservative 139; Mismatches 409; Indels 237; Gaps 36;

QY 94 RORRAAGGI---LHLELLVAVGPDVFOAH-QEDTERRYVLTNLTNGAELLRDPDLGAQPRV 149
Db 281 RPRRSAGKQKGLNVELLVADKQWEKHGKGNVTYLLTVMVMSGLFKDGTIGSDINV 340

QY 150 HLVRKVLTEPEGAPNITANLTSLLSVCGMSQNTINPEDDDTPGHADLVLYITRFDL-EL 208
Db 341 VVYSLILLEOPEGLLINHEADQSLNSFCQWSALIGKNGKRHDHA---ILLTGFIDCSW 397

QY 209 PDGNRQVRGVTVQLGGACSPWTSCLITEDTGDLGVTTIAHEIGHSGFLEHGDGPGSGCGPS 268
Db 398 KNEPCDTLGFAPISGMCKSYRSCNTINEDTGLGFTAIAHSGHNGFMTHDG-EGNPCRKA 456

QY 269 -GHVMSDGAAPRAGLAWSPCSRRLSLLSAGRARCVMDPRPPOGSGAGHPDPAQPLGY 327
Db 457 EGNIMPTLTGNNGVFSWSSCSRQYLKFLSTPQAGCLVDPEK-QAGQYKY-PDKLPQOI 514

QY 328 YSANECQVAFGPKAVACT--FAREHLDMCOALSCHTDPLDQSSCSRLLVPLLLDGTCEGV 385
Db 515 YDADTQCKWQFGAKLCSLGFVK--DICKSLWCHR---VCHRCETKFMFAAEGTVCOL 568

QY 386 EKWCCKGRCSRSLVELTPIAAVHGHSWGPSPCSRSCGGGVVTRRRQCNPRPFAFGRA 445
Db 569 SMCRCQGCQVKFGBELGP-RPIHGQWSAWSKWSKSECSRTCGGVKFOERHCNPKPQVGGIF 627

QY 446 CVCADLQAEACNTQACBKTQLEFMSQCCARTDGPRLRSSPGGASFYHGAAPVPHSQDAL 505
Db 628 CPSSRIYQLCNINPCNENSLDFRAQCAEYNSKPF-----GWFIQWKPYTKVVEEDR- 681

QY 506 CRHMCRAIGBSFIMKRGDSFLDGTFRCPSPGPRDGTLSLVCVSGSCRTFGCDGRMDSQQWY 565
Db 682 CKLYCKAENETFFFMAGSKVKGDTGPCSPN--KND-----VCIDGVCELVCCHDELGSKAVS 735

QY 566 DRQVCQGDNDSTCSPRKGSP-TAGRAREYVTLTVPNTLSVYIANHRPLFTHLAVR-IG 623
Db 736 DAGVCVKGDNDSTCKFYKGLYNQKANEYYPWIIIPAGARSIEIQELQVSSSYLAVRSL 795

QY 624 GRVYVAGKMSIPNTYPSLLEDGRVEYRVALTDRPLRLEBIRIWGPQEDADIQVYRR 683

Db 796 QKYLATGGSIDWPGEPP--FAGTTFEYQSRFN-----RPERLYAPGPTNETLVPEILMQ 848
QY 684 YGEYGNLTRPDITFTYFQPK-----PRQAWVAARVGPSCVSCGAGLRWVNSCL 734
Db 849 -GK-----NPGIAWKVALPKVMNGTPPATKRPAYTWSIVQSECSVSCGGYINVAICL 901

QY 735 DQARKELVEIVQCGSOQPPAWPEACVLEPCPPYKAVGDFGPGCSASCGGGLRSPVRCVYE 794
Db 902 RDQNTQ-VNSSFCSAKTKPVTEPKICNAFSCPAVMPGSEWSTCSKACAGGQSRKICQCVQ 960

QY 795 AQGSLKLTLPARCRAGAQAQPAVALETCPQPCPARWEVSPSPSCTTSAGGAGLALENETC 854
Db 961 -----KKPFQKEEAVLHSLCP-----VSTPTQV----- 983

QY 855 VPGADGLEAPVTEGPGSVDEKLPAPEPCVGMSCPPGWHGLDATSAGEKAPSPWGSRTGA 914
Db 984 -----QACNSHACPQW-----SLG-----PW----- 1000

QY 915 QAAHVWTPAAGSCSVSCGRGLMELRFLCMDSALRVFVQEBELCGLASKPGSRREVCQAVPC 974
Db 1001 -----SQSKTCGRGVKRELLCKGSA----- 1022

QY 975 PARWQYKLAACSVSCGRGVVRRILYCARAHGEDDGEIILLDTQCQGLRPEPEACSLPE 1034
Db 1023 -----AETLPESQCTSLPRPELQEGCVLGR 1047

QY 1035 CPP-----RWKMSILGPCSASGLGTARRSVACVOLD-OGQDVEVDEAAACAALVRPEASVP 1089
Db 1048 CPKNSRLQWVASSWSECSATCGLGVKREMKCSKFGKGLITFFERRCENIKKPNLD-- 1105

QY 1090 CLTADCTYR-----MHVGTWMECSVSCGDIQRRRDTCLGFPQAQAPVADFC 1136
Db 1106 -LEETCNRRACPAHPVYVMVAGWYSLPWQCTVTCGGVQVRESVHCV-----CQGRSSSC 1160

QY 1137 QHLPKPVTVRGCGWAGPC 1153
Db 1161 LLHQKPPVLRACTNFC 1177

RESULT 13
US-09-972-467-2
; Sequence 2, Application US/09972467
; Patent No. US20020090373A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; FILE OF INVENTION: USES THEREOF
; FILE REFERENCE: PCT0850A
; CURRENT APPLICATION NUMBER: US/09/972,467
; CURRENT FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1629
; TYPE: PRT
; ORGANISM: Human
US-09-972-467-2

Query Match 15.2%; Score 1176.5; DB 9; Length 1629;
Best Local Similarity 25.5%; Pred No. 1.7e-77;
Matches 348; Conservative 170; Mismatches 505; Indels 343; Gaps 50;

QY 89 QQRQRORRAAGGILHLELLVAVGDFVFOAHQEDTERRYVLTNLTNGAELLRDPDLGAQPR 148
Db 279 KRTHRRTRKFLSYPRFVEVLVADNRVSVYHGENLQHYILLMSIVASYKDPISGNLIN 338

QY 149 VHLVKVILTEPEGAPNITANLTSLLSVCGMSQNTINPEDDDTPG--HADLVLYITRFDL 206
Db 339 IIVVNLIVHNEQDGPISFNAQTLLKNFCQWQHSKN-----SPGGIHHDTAVLLTRQDI 393

QY 207 ELFDGNRQVRGVTVQLGGACSPWTSCLITEDTGFDLVGTIAHEIGHSGFLEHGDGPGSG 266
Db 394 CRAHDKCDTLGLAELGTICDPYRSCSISDLSGLSTAFTHAELGHVFNPHD--DNNKCK 451

QY 1162 -----VPHEAAAPGRTTATPAGACGR-----CHLE-----PTG 1190
Db 1431 REQCNTHAAPHDAWSTGFWSSCV-SCGRGHKORNYVYCMADGSHLESYCKHLAKPHG 1489
QY 1191 TIDMRG-----PGQADCAVAIGRPLGEVVTILVLESSLNCASG 1228
Db 1490 HRKCRGRCFCFKWAGAWSQCSVSCGR-----GVQORHVGCQIG 1527

RESULT 14
US-10-274-639-10
; Sequence 10, Application US/10274639
; Publication No. US20030232349A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: DELEGANE, Angelo M.; GANDHI, Ameena R.
; APPLICANT: HAFALIA, April J.A.; LU, Dying Aina M.
; APPLICANT: PATTERSON, Chandra; TRIBOULEY, Catherine M.
; APPLICANT: DAS, Debopriya; KALLICK, Deborah A.
; APPLICANT: NGUYEN, Dannel B.; LEE, Ernestine A.
; APPLICANT: KEAN, Farrah A.; YUE, Henry
; APPLICANT: AU-YOUNG, Janice K.; GRIFFIN, Jennifer A.
; APPLICANT: POLICKY, Jennifer L.; RAMKUMAR, Jayalaxmi
; APPLICANT: YANG, Junning; THANGAVELU, Kavitha
; APPLICANT: DING, Li; KEARNEY, Liam
; APPLICANT: BAUGHN, Mariah R.; BOROWSKY, Mark L.
; APPLICANT: SANJANWALA, Madhusudan M.; YAO, Monique G.
; APPLICANT: BURFORD, Neil; WALIA, Narinder K.
; APPLICANT: LAL, Preeti G.; LEE, Sally
; APPLICANT: TODD, Stephen; LO, Terence P.
; APPLICANT: TANG, Y. Tom; ELLIOTT, Vicki S.
; APPLICANT: AZIMZAI, Valda; LU, Yan
; TITLE OF INVENTION: PROTEASES
; FILE REFERENCE: PI-0167 USA
; CURRENT APPLICATION NUMBER: US/10/274,639
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: PCT/US01/22397
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/220,063
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/221,680
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: US 60/223,544
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/224,717
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/225,988
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 60/227,568
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1916
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030232349A1 2994162CDI
US-10-274-639-10

Query Match 15.1%; Score 1172.5; DB 15; Length 1916;
Best Local Similarity 25.4%; Fred. No. 4.2e-77; Indels 343; Gaps 50;
Matches 347; Conservative 171; Mismatches 505;
QY 89 QRCQRCORRAAGGILHLELVAVGPDVFOAQHEDTERTYVLTNLINIGALLRDPFSLGQAFR 148
Db 260 KRTHRRTRKFLSYPRFVFLVADNRMVSYHGENLOHYLTLMMSIVASYKDSIGNLIN 319
QY 149 VHLVKRWILTPEGAAPNITANLTSSLLSVCGSQTNPEDDTDPG--HADLVYITFDL 206
Db 320 IIVNLVLIHNEQDQGSISFNAQTLLKNFCQWQHSKN-----SPGGTHHDTAVLLTRQDI 374

QY 267 PSG-----HYMASDGAAPRAGLAWSPGSRRLQLLSLSAGRARCVWDPPRPOFGSAGHP-P 320
Db 452 EEGVKSPQHVAPTLNFYTNMWSKSKRYITEFLDTGYGECLINEPESRP-----YPLP 507
QY 321 DAQPGLYYSANEORVAFGKAVACTARHLDLWCOALSCHTDPLDQSSCSRLVPLLDG 380
Db 508 VQLPGILYNNKQCELLIFGPGSQVCPYMQ-----CRRLLWCNNVGVHKGCRTOHTPWADG 563
QY 381 TEGCEVKWCSKGRCSLVELTPIAAVGRWSSWSPRSPCSRSCTGGGVVTRRQCNNRPPA 440
Db 564 TCEBPKCHKYGF- -VPKENDVPTVDTGWSGWSNFFGTCSRTCCGGIKTAIACNRPPEPK 621
QY 441 FGGACVGDLOABMNCNTQACEKTLQEFMSQCCARTDQOLRSSPGGASFYHMGAAVPHS 500
Db 622 NGKYCVGRMFKSCNTEPCLKQKQRPDRBQCAHFQKGFEN-GLLPNVRW---VPKY 677
QY 501 QGDAL-----CHRMCAIGESFMKSGDFLDTGRCMPSPREDGTLSLVSGSCRTFGCDG 557
Db 678 SGILMKDRCKLFCRVAGNTAYQLRDRVIDTGPC-----GQDTNDICVQGLCROAGCDH 731
QY 558 RMDSQWQDQVCGGDNSTCSPRKSGSTAGAREYVTLVTPNLTSVYIANHRPLFT- 616
Db 732 VLNSKAARDKCGVCGDNSSCKTVAGTENTVH-YGYNTVTRIPAGATNIDVRQHS--FSG 788
QY 617 -----HLAVR-----IGRYVVA-GKMSISPNNTYPSLLEDGRVEYRVALTEDRLPR 562
Db 789 ETDDNNVLALSSSGEFLNGVFWTAKRIR-----IGNAVEVSGSET-----A 835
QY 663 LEEIRINGPQOEDADIOYRRYGEYGNLTDPDITFTY---FQPKRQAWV----- 711
Db 836 VERINSTRIEQLLEQLVL-----SVGLYNPDVRYSNFIPIEDKPOQ-FYWNHGFQMA 889
QY 712 -----NAVRGPCSV 720
Db 890 CSKPCQGERKRLVCTRESQDQTVSDQCDRLPOFGHITEPCGTDCLRWHVASRSECSA 949
QY 721 SCGAGLEWVN-----YSLDQARKELVETVQCGSQQPPAMPEACVLEPCPPVAVGDG 775
Db 950 QCGLYGTLDLYCAKYSRLD-GKTEKVDGFC-SSHPKPSNREKSGEBCNTGWRYSAT 1007
QY 776 PCSASCGGGLRERVRVCAQSLKTLPPARCAGAAQOFAVALETNCPQPCPARVESE 835
Db 1008 ECKSCDGGTQRRRAICVNTENDV---LDDSK---THQKVTIQRCSFPCP-QWKS 1060
QY 836 PSCTSGAGGAGLALENETCFVAGDGLAPVTEGPGSVDBKLPAPFCVQMSCP-----PGW 891
Db 1061 WSECLVTCGKHQRQVWCQFGEDRLNDRMCD-----PETKPTSMQTCCOPECASW 1111
QY 892 GHLDATGAGEKAPFPWGSIRTGAAAHVWTPAAGSCSVSCGRGLMELRFLCWDLSALRPV 951
Db 1112 -----QAG-----FW-----GQCSVTGQGYQLRAVKKCIIGTYMSVV 1143
QY 952 QEEELCGLASKPGSREVCQAVPC---PA-----RWQY-KLAACSVSCRGV 993
Db 1144 DDNDNATRTDTOD-CELPSCHPPAPAPETRTSTYSAPRTQWRFGSWTPCSATCGKT 1202
QY 994 VRRILYCARAHGEDDGEILLDTQCGLPRPBPQACSLPECPPRRWKMSLGPSCSACGL 1053
Db 1203 RMRVSC-----RDENGVADESACATLPRVAKESCVTPC-GQNKALDWSSCSVTCGQ 1256
QY 1054 GTARRSVACVQLDQGDVEVDBAACALVRPEASVPCLIADCTYR----- 1098
Db 1257 GRATRQVMCVNY---SDHVDRSECDQDIYIPETDQDCSMSPCQRTPDGSLAGHPFQNE 1313
QY 1099 -----WHVGTWECVSVCGDGIQRERDTCGLGPOAAQVVPADFQHL 1140
Db 1314 YRPSASPSRTHVLGGNQWRTPGACSTTCAGGSQRRVVVC---QDENGTYANDCVERI 1370
QY 1141 KPVTVRGCVAGPCV-----GQGTPEL----- 1161
Db 1371 KPDEORACESGPCOWAYGNWGECKLCCGGITRLVVCORSNGERFPDLSCBILDKPPD 1430

```
QY 207 ELPDGNROVGVTLQGGACSPWNSCLITDGTDFDLGVTHIAEIGHSGFLEHGDGPGSGCG 266
Db 375 CRAHDKCDTLGLAELGTICDPYRSCSISDGLSTAFTHAELGHVFNPHD--DNKKCK 432
QY 267 PSG-----HYMASDGAAPRAGLAWSPCSRRLQLLSLAGRARCVMPPRPOPSAGHP-P 320
Db 433 BEGVKSPQHWAFTLNFTYTPMWSKSKRYITFELDTGYGECLENEPSRP-----YPLP 488
QY 321 DAQPGLYYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLLVPLLDG 380
Db 489 VOLPGLILYNNVKQCELIFGPGSQVCPYMMQ-----CRRLLWCNNVNGVHKGCRTOHTPWADG 544
QY 381 TECGEVKWCKGRCSRLVELTPIAAVHGWSWSPRSCSRSCGGGVVTRRQCNPRPA 440
Db 545 TECEPGKHCKYGC--VPKEMDVPVTDGWSGWSWSPFGTCSTRCTGGIKTAIRECNRPEPK 602
QY 441 FGRACVAGDLQAEQCMNTQACEKQLEPMQOCARTDQPLRSSPGGASFYHWAAGVPHS 500
Db 603 NGKVCVGRNKKFKSCNTEPCLKQKRFDRDQCAHFHDKHFNIN-GLLPNVRW---VPKY 558
QY 501 QGDAL----CRHMCRAIGESFIMKGGDSFLDGTFRCPMSGPREDTGLSLCVSGSCRTFGCDG 557
Db 659 SGILMKDRCKLFCRVAGNTAYQLRDRVIDGTFC-----GQTDNDICVOGLCROAGCDH 712
QY 558 RMDSQVMDRCQVCGDNSTCSPKSGFTAGRAREYVTLVTPNLTSVYIANHRLPFT- 616
Db 713 VLNSKARRDKGVCVCGDNSSCKTVAGTENTVH-YGYNTVRIIPAGATNIDVRQHS--FSG 769
QY 617 -----HLAVR-----IGRYVVA-GKMSISPNITYPSLLEDGVEYRVALTEDRLR 662
Db 770 ETDDNYLALSSSGEFLNGFVVTMAKREIR-----IGNAVVEYSGSET-----A 816
QY 663 LEBIRIWPQEDADIQVRYRBYGNLTRDITFTY---FQPKPQAAWV----- 711
Db 817 VERINSTRIEQELLQVL-----SVGKLNDVRYISFNIPIEDKQD-FYNSHGFPWA 870
QY 712 ----- 720
Db 871 CSKPCQGERKELVCTRESQTLVSDQRCDRLPQPHITEPCGTCDCLRWHVASRSECSA 930
QY 721 SCGAGLRWN-----YSCLDQARKELVETVQCGSQQPPAWPEACVLEPCPPYVAVGDFG 775
Db 931 QCGLGYRTLDIYCAKYSRLD-GKTEKVDGDFC-SSHPKPSNEKCSGECNCTGGRYSAWT 988
QY 776 PCSASCGLLEERPVRCVEAQSLKLTLPAPCRAGAQAQPAVALETNCPQCPARVEUSE 835
Db 989 ECSKSCDGTORRAICVNTNDV---LDDSKC---THQEKVTIQCSEFFCP-QWKSQD 1041
QY 836 PSSCTSAGGAGLALENETCVPGADGLEAPVTEGPGSVDEKLPAPBPCVGMSCP---PGW 891
Db 1042 WSECLVTGKGKHKRWQVQCGEDRLNRMCD-----PETKETSMTQCOQPECASW 1092
QY 892 GHLDATSAGEKAPGSPWGSIRITGAQAHAHVMTPAAGSCSVSCGRGLMELRFLCMDSALRPV 951
Db 1093 -----QAG-----PW-----GQCSVTGGQYQLRAVKCIIGTYMSV 1124
QY 952 QEBELCGLASKPGSRREVCOAVPC---PA-----RMQY-KLAACSVSCRGV 993
Db 1125 DDNDCAATRPDTQD-CELPSCHPPPAAPATERRSTYSAPRTQWRFGSWTPCSCATCKGT 1183
QY 994 VRRILYCARAGEDGGEILLDTQCGLPRPPEOEACSLPECPPRKMWLSLGCSCSACL 1053
Db 1184 RMRYVSC-----RDENGSADESACATLPRPAKECSVTPC-GQWKALDWSSCSVTGQ 1237
QY 1054 GTASRSVACVQLDQDQVEVDAAACALVRPEASVPLIADCTYR----- 1098
Db 1238 GRATQVNCVNY---SDHVDRSECDQDQYIPKTDQDCMSPCQPTPDSGLAQHPQONED 1294
QY 1099 -----WHVGTWMECSVSGDGIQRRRDTCLGPOAQAPVPADFCQHL 1140
Db 1295 YRPRSASPSRTHVLOGNQWRTGFWGACSTCAGGSQRRVVC---QDENGYTANDCVERI 1351
QY 1141 KPVTVRGCAWGPV-----GQGTPLS----- 1161
```

Db 1352 KPDEQRACSGPFPQWAYGNWGBCTKLCGGGIRTELVVQCQRNGRFPDLSCEILDKPPD 1411

QY 1162 -----VPHEEAAAAGRITATPAGACGR-----OHLE-----PTG 1190

Db 1412 REQCNTHACPDAAWSTGPMSSCSV-SCGRGHKQRNVYCMAXDGHLESYDCHSLAKPHG 1470

QY 1191 TIDMRG-----PGQADCAVAIGRPLGEVVVTLVLESSLNCSAG 1228

Db 1471 HRKCRGRCPCPKWAGAWSQCSVSCGR-----GVQQRHVGCQIG 1508

RESULT 15

US-09-918-171A-13

; Sequence 13, Application US/09918171A

; Patent No. US20020110894A1

; GENERAL INFORMATION:

; APPLICANT: Apte, Suneel

; APPLICANT: Hurskainen, Tiina L.

; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases

; FILE REFERENCE: 26473/04193

; CURRENT APPLICATION NUMBER: US/09/918,171A

; CURRENT FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 09/369,364

; PRIOR FILING DATE: 1999-08-06

; NUMBER OF SEQ ID NOS: 31

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 13

; LENGTH: 1882

; TYPE: PRT

; ORGANISM: Homo sapiens ADAMTS-9

; FEATURE:

; NAME/KEY: MOD.RES

; LOCATION: (468)

; OTHER INFORMATION: Xaa = Cys

; NAME/KEY: MOD.RES

; LOCATION: (521)

; OTHER INFORMATION: Xaa = Tyr

US-09-918-171A-13

Query Match 14.9%; Score 1157; DB 9; Length 1882;

Best Local Similarity 25.5%; Pred. No. 5.7e-76;

Matches 348; Conservative 170; Mismatches 505; Indels 342; Gaps 52;

QY 89 ORQORQORRAAGGILHLELLVAVGPDVQAHQEDERYVLTNINIGALLRDRSLGQAF 148

Db 227 KRTTRTKRFLSPRPFVEVLVADNRVMSYHGENLQHYLTLLMSIVASYIKDPSIGNLN 286

QY 149 VHLVKNVILTEPEGAPNITANLTSLSSVCGWSQTINPDDTDPG--HADLVLYITRFDL 206

Db 287 IVIVNLIVHNEQDGFISFNAQTILKNFCQW-----QHSNPSGGIHHDTAVLLTEQDI 340

QY 207 ELPDGNROVGVTLQGGACSPWNSCLITDGTDFDLGVTHIAEIGHSGFLEHGDGPGSGCG 266

Db 341 CRAHDKCDTLGLAELGTICDPYRSCSISDGLSTAFTHAELGHVFNPHD--DNKKCK 398

QY 267 PSG-----HYMASDGAAPRAGLAWSPCSRRLQLLSLAGRARCVMPPRPOPSAGHP-P 320

Db 399 BEGVKSPQHWAFTLNFTYTPMWSKSKRYITFELDTGYGECLENEPSRP-----YPLP 454

QY 321 DAQPGLYYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLLVPLLDG 380

Db 455 VOLPGLILYNNVKQCELIFGPGSQVCPYMMQ-----CRRLLWCNNVNGVHKGCRTOHTPWADG 510

QY 381 TECGEVKWCKGRCSRLVELTPIAAVHGWSWSPRSCSRSCGGGVVTRRQCNPRPA 440

Db 511 TECEPGKHCKYGC--VPKEMDVPVTDGWSGWSWSPFGTCSTRCTGGIKTAIRECNRPEPK 568

QY 441 FGRACVAGDLQAEQCMNTQACEKQLEPMQOCARTDQPLRSSPGGASFYHWAAGVPHS 500

Db 569 NGKVCVGRNKKFKSCNTEPCLKQKRFDRDQCAHFHDKHFNIN-GLLPNVRW---VPKY 624

[illegible]

Search completed: March 13, 2004, 08:07:02
Job time : 119.1 secs

381 TEGVEKWCGRCELSVELTPIAAVHGWSWGPSPCSRSCTGGVVTERRRQNNRPPA 440
511 TEGPFGKCKXGFC--VPKEMDVFTDGSWGSWSPFGTCSRCTGGGKTAIRECNREPK 568
441 FGGRACVADLQAEKCTQLEFMSQOCARTDQGLRSPGSGASFYHGAAPVHS 500
569 NGKYCVGRMKFKSCNTEPCLKQKDFDEQCAHFDGKHFNIN--GLLPNVRM--VPKY 624
501 QGDAL---CRHMCRAIGSFYKWRGDSFLDGTCTCMSPGPREDTLSLCVSGSCTFCG 557
625 SGILMKDRCKLFCRVAGNTAYQLRDVDTGTPC-----GQDTNDICVQGLCRQACDH 678
558 RMDSQOVDRCOVCGDINSTCSPKSGFTAGRAREYVTLTVPNTSVVIANHRPLFT- 616
679 VLNSKARDKCGVCGDNDSCKTAGTNTVH--YGNVTVVRIPAGATNIDVROHS--FSG 735
617 -----HLAVR-----IGGRYVVA--GKMSISPNTTYPSSLDDGRVEYRVALTEDRLPR 662
736 BTDDNYLALSSKGEFLNGNFVVTMAKREIR-----IGNAVVEYSGSET-----A 782
663 LEEIRINGPLODADIQVYRYGEEYGNLTPDITTY-----FQPKPQAM-----VMAAV 714
783 VERINSTRIEQELLQVL-----SVGKLYNPVDRYSFNIPEDKPOQFYWNHSGPQWAC 837
715 RGPC-----SVS-----CGAG--LRW----- 728
838 SKPCQGERKRLVCTRESQDLTVSDQCDRLPQFGHITEPCGTCDLRHWVASRSECSAQ 897
729 -----VNSCLDQARKELVETVQCQSQQPAMPEACVLPCCPYMAVGDFGP 776
898 CGLGYRTLDIYCAKYSRLD--GKTEKVDGDFC--SSHPKPSNREKCSGBCNTGGWYSWTE 955
777 CSACGGGLRERPVRCVEAQSGLKLPARCRAGAAQPAVALETCPNQPAPARWEVSEP 836
956 CSKSCDGGTORRAICVNTDNDV--LDDSK--THQEKVTIQRCSFPCP--QWKSQDW 1008
837 SSCTSAGGAGLALNETCPVAGDLEAPVTEGPGSVDEKLPAPPCVGMSCP-----PWG 892
1009 SECLVTCGKHGKHSQVWCQFGEEDLNDRMCD-----PETKPTSMQTCQPECASW- 1058
893 HLDATSAGBKARSPWGSIRTAQAHAHVWTPAASCSVSCGELMELRFLCNDLSALRVPVQ 952
1059 -----QAG-----PW-----VQCSVTCCGGYQLRAVKCIIGTMSYVD 1091
953 BELCGLASKGSRREYCOAVPC---PA-----RWOY--KLAACSVCGRGVV 994
1092 DNDCAATRPTDQD--CELPSCHPPPAAPETRRSTYSAPRTQWRFGSWTPCSATCGKTR 1150
995 RRLYCARAHGEDDGBEILLDTQCGLPPEPOBACSLPECPPRKWKVMSLGPSCSACGLG 1054
1151 MEYVSC-----RDENGVADESACATLPRPVAKBECSTVPC--GQWKALDWSGCVTCGQ 1204
1055 TARRSVACVOLDOGVDEVEDEACALVRPEASVPLIADCTYR----- 1098
1205 RATRQWCVNY---SDHVIDRSECDQYIPETDQDCMSPCPQRTDPSGLAQHPQNE 1261
1099 -----WHVGTWMECSVCGDGTORRRTDCLGPAQAPVPADFCQHLPK 1141
1262 RPRASPSRTHVLGNQWRTGPMGACSTCAGSQRRVVC---QDENGXTANDCVERIK 1318
1142 PVTVRSCWAGPCV-----GOTPSL----- 1161
1319 PDEQACESGCPQWAGNNGECTKLCGGGIRTLVVSQRNGERFPDLSCILLDKPDR 1378
1162 -----VPHEEAAAPOGRTTATPAGACGR-----QHLE-----PTGT 1191
1379 EQCNTHACHADAAWSTGFWSSCV--SGRGHKQRNVYCMAXDGSHLESDYCKHLAKPHG 1437
1192 IDMRG-----PGQADCAVAIGRPLGEVVVTLRVLESLSNCAG 1228
1438 RKGRCGRCPKWKAGAWQCSVSMGR-----GVQQRHVGCQIG 1474

RESULT 2
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the s
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match 14.8%; Score 1148.5; DB 4; Length 1224;
Best Local Similarity 28.3%; Pred. No. 2.3e-84;
Matches 328; Conservative 145; Mismatches 443; Indels 241; Gaps 43;

QY 2 GRTPRPKAPSHAPILGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWSPSHFQQSCL 61
DB 192 GRAAGSSPSHV-----LYKSTEPHAP-----GASEVLVTSTWELAH----- 230
QY 62 QALEPQAVSSYLSPGAPLK---GR-----PPSPG-----FQRCQRRRAA 99
DB 231 QPLH-----SSDLRLGLPQKHFCGRKKRYMPQPKEDLFLPDEYKSLRHKRSLRSHR 286
QY 100 GGIHLELVAVGPDVFOAH--QEDTERYVLTNLNIGAEILLRDPGLSGAOFRVHLVKWVILT 158
DB 287 NEELNVEITLVVDKWMQNHENITTVTLTNMVSALFKDGTIGGNINIAVLILLE 346
QY 159 BEGAPNITANLTSSLLSVCGWSQTNPEDTDPGHADLV--LYITFPDLPLDGNRQVR 216
DB 347 DEQPGLVISHADHTLSSFCQWQSLMGKDGTRHDHAILLTGLDICSWKNEPCD---TL 402
QY 217 GVTQLGGACSPWSCILITEDTGFDLGVITAEHIGHSFGLHGDGAPGSGCGPS--GHVMA 275
DB 403 GFAP--SGMCKSKYRSCINEDTGLGTAHESGHNFMIHDG--EGNCKKSEGNIMSP 461
QY 276 GAAPRAGLAWSPCSRQLLSLISAGRACVYMDPPRQPGSAGHPDPDAGFLYYSANEQCR 335
DB 462 LAGNCGVFWSPCSRQYHKLSTAQAICLAD--QPKPVKEYKYPEKLPGLYDANTOCK 519
QY 336 VAFPGKAVACTFAREHLMQALSCHTDPLDQSCSRLVPLLDGTGCGVEKWCCKGR 395
DB 520 WQFGEKAKLCMLDFKK--DICKALWCHR---IGRKCETKFMPEABGTICGHDWMCRCGQCV 575
QY 396 SILVELTPIAAVHGWSWGPSPCSRSCTGGVVTERRRQNNRPAFGGRACVADLQAE 455
DB 576 KYGDEGP--KPTHGHWSDWSSWSPCSRTCGGVSHRSRLCTNPKPSHGKFCGEGSTRTL 634
QY 456 CNTQACEKTLFMSQOCARTDQGLRSPGSGASFYHGAAPVHSQ---GDALCRHMCRAI 513
DB 635 CNSQKCPRSDVFPAAQCAEHNSRRFR-----GRHYK---KPYTQVEDQDLCKLYCIAE 686
QY 514 GESFIMKRGDSFLDGTCTCMSPGPREDTLSLCVSGSCTFCGDRMDSQQVMDRCQVCG 573
DB 687 GDFPFSLSNKVKDGTFC-----SEDSRNVICDIGICERVGCDNVLGSDAVEDVCGVCG 740
QY 574 DNSTCSPKSGFTA--GRAREYVTLTVPNTSVVIANHRPLFTHLAVR-----IGR 625
DB 741 NNSACTIHRGLYTKHHHTNOYHYMVTIPSGARSIRIYEMAVSTSVISVRNALLRRYINGH 800
QY 626 YVY--AGKMSISNTTYPSSLEDRVEYRVALTEDRLPRLEIRIWGPLEQADAIQVYR 683
DB 801 WTVDWPGRYKFS--GTFP-----DYRSYNEP-----ENLIATGPTNETLIVELLFQ 845

Qy 684 YGEBYGNLTPDITTFY-----QPKPQAWMAAVRPGSVSCGAGLRWVNYSCLD 735
Db 846 G-----RNPVAMEYSMPRLGTEKQPPAQPSTWAIVRSECSVSCGGQMTVREGCYR 898
Qy 736 QARKELVETQCGSQOPANPEACVLEPCPPYWAAGDFPCSNASCGGLRERVRVCVEA 795
Db 899 DLKFO-VNMSFCNPKTRPTVGLVPCVSAFCPPSWSVGNWSACSRTCCGGAQSRFVQCTRR 957
Qy 796 QGSLKLTTPARCRAGAQAVALTECNPOPCPARWEVSPSSCTSGAGGALALENETCV 855
Db 958 VHYDEPVPASLC-----POPAS----- 976
Qy 856 PGADGLEAPVTEGPGSVDEKLPAPCEVGVNSCPFGWGHLDATSAGEKAPSPWGSIRTAQA 915
Db 977 -----SRQACNSQSCPPAW-----SAG-----PW----- 995
Qy 916 AAHVTPAAGSVSCVSCGRGLMEIRFLCMD---SALRVPVQEBELCGLASKPGSRREVQAV 972
Db 996 -----AECSTCGKGRKRAVACKSTNPSARAQLLPDAVCTSPKP-RMHEACILQ 1045
Qy 973 PC-PARWQYKLA---CSVSCRGVVRILYCARAHGEDDGEIILLDTQCGLRP--E 1025
Db 1046 RCHKPKKLOLVSAWSQCSYTCERGTOKRFKCAEKVSGKYRE-LASKKSHUPKPSLE 1104
Qy 1026 POEACSLPCPPRWKVMNLGP-----CSASCGLGTARRSVACVOLDOQDVEV 1073
Db 1105 LERACAPLCPRHPPFAAGPSGWSFASPWSQCTASCQGVQTRSVQC--LAGRPA-- 1160
Qy 1074 DEAAACAALVRPEASVPC 1090
Db 1161 --SGCLLHQKPSASLAC 1175
RESULT 3
US-09-800-729-155
; Sequence 155, Application US/09800729
; Patent No. 6805592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 155
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-155
Query Match 14.4%; Score 1118; DB 4; Length 2165;
Best Local Similarity 26.6%; Pred. No. 1.6e-81;
Matches 331; Conservative 142; Mismatches 501; Indels 272; Gaps 42;
Qy 91 QROQRRAAGILH-LELLVAVGPDVFOAHQEDTERTYVLTNIGAEILRDFSLGQFRV 149
Db 267 RKARSRAANSWDHYVELVVDATWQVEYHGRSLEDYVLTFLFSTVASIYRHQSRLASINV 326
Qy 150 HLKVMVLTTEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRPLELP 209
Db 327 VVVKLIIVLKTENAGPRITQNAQQTLDQFCRWQYQYNDPDDSSVQHHDVAILTRKDICKS 386
Qy 210 DGNRQVGVTLGAGCSPWSCLITEDTGFOLGVITAEIGHGSFGLHDCGAPSGCGCP-- 267
Db 387 QGKCDTGLAELGTWCDMOKSCAILEDNGLSAFTIAELGHVFSIPHDDE--RKSTYM 444
Qy 268 -----SGHVNASDGAAPRAGLWSPCSRRLQLLSL--SAGRARC 305

Db 445 PVNKVCKPQSTFKDTQFQNNFHIMAPTLEYNTHPWSWSPCSAGMLRFLFENRGTOCL 504
Qy 306 WDFPRPQPGSAGHPDDA---QFGLYISANEQCRVAFPGKAVACTFAREHLDMOALSCH 361
Db 505 FD---QPVERRYEDVFRDBEGKYDAHQCKFVFGPASELCFY---MPTCRRLMCA 556
Qy 362 TDPLOQSSCSRLVLLDGTGCEVKE--WCSKGRCSRLVELTP--IAAVHGRWSSWGPRS 417
Db 557 TFYGSQMGCRTOHMPWADGTPCDESRSFCHHGAC---VRLAPESLTAKIDQWGDWRSG 613
Qy 418 PCSRS CGGVVTRRRCONNPRPAFGGRACVAGDADLOAEMCNTOACEKTQLEFMSQOCARD 477
Db 614 ECSRTCGGVQGLRDCDSPKPRNGKCYVQBRERYSCNTQCPWDTPQPREVOCSEFN 673
Qy 478 GQPLRSSPGASFYHMGAAVPHSQGDALCRHMCRAIGESFTMKRGDSFLDGTGRCMESGR 537
Db 674 NKDIGIQGVASTNTHVVPKYANVAPNERCKLYCRLSGSAAFYLLRDKVVDGTPCDRNGD- 732
Qy 538 EDGTLVSVSGS CRTFCGDRMDSQQWDRVCQVCGGNSTCSPRKGSSTAGRARVYVFL 597
Db 733 ----DICVAGACMPAGCDHQLHSTLRDKCGVCGGDDSSCKVVKGTENEQGTFGYNEVM 787
Qy 598 TVTPNLTSVIIA----NHRPLFTHLAVR-IGGRYVAVAGKMSISNTTYPSSLLEDGRVEY 651
Db 788 KIPAGSANIDIRQGYNNMKEDDNYLSRAANGEFLLNGHFQVS-LARQQIAFQDVTLEY 846
Qy 652 RVALTEDRLPRLEIRIWFGLQEDADIQVYRERYGEEVGNLTPRDPITFTY-----FQ 702
Db 847 S---GSDAI--IBRINGTGPIRSIYVHVL-----SVGS-HPPDISIYEMTAAPNAVIR 895
Qy 703 P-----WVNYIS-----CLDQARKELVETVQCCSQSQQPAPWPAEACV 761
Db 896 PISSALYLMWVTDWTWTECDRACRQSQSKLCLMDSTHQRSHDRNCQNVLPKQATMCN 955
Qy 709 -----WVMAAVRPGSVSCGAGLR----- 727
Db 956 IDCSTRWITEDV--SSCSAKCGSGQKQKRVSCVMGEGDRQTPASEHLCDRNSKPSDIASCY 1014
Qy 728 -----WVNYIS-----CLDQARKELVETVQCCSQSQQPAPWPAEACV 761
Db 1015 IDCSGRKNYGEWTSCTSETCGSNGKMKRKSVCVDSNRRVDESL-C-GREQKEATRECN 1072
Qy 762 LEPCPPYVAVGDFGPCSACGGGLRERPVRCVEAAGSLLKTLPPARCAGAAQPAVALET 821
Db 1073 RIFC-PRWVYGHVSECSRSCDGGVAMHAQCLDAADRETHT---SRC-----GPAQTOEH 1123
Qy 822 CNPQPCPARWEVSEPPSCTSGAGGALALENETCPVGDGLBAPVTEGPGSVDEKLPAPPEP 881
Db 1124 CNEHAC-TWQFGVWSDCSAKCGDGVQYRDANC---TDHRSVLPHEHRCLEMEKI-ITKP 1178
Qy 882 CVGMSCPPGWGHLDATSAGEKAPSPWGSIRTAQAQAAHVWTPAAGSCSVSCGGLMELRFL 941
Db 1179 CHRESCF-----KYLGE-----WS-----QCSVSCEDGWSRRVS 1209
Qy 942 CMPSALRVVQEBELCGLASKPGSRREVQAVPCPARWQYKLAACSVSCRGVVRRLYCA 1001
Db 1210 CV-SGNGTEVDMSLCAGTASDRPASHQTNLGTCPFRWNTDWSACSVSCGIGHRETECI 1268
Qy 1002 RAHGEDDGEIILLDTQCGLRPEPFOEACSLPEPCPPRWKVMWSLGPSCSACGLGTARRVA 1061
Db 1269 YRQSVDA-----SFGDITKMPETSQTCHELLPC-TSKPSSHWSFCSVTCGSGIQTRSVS 1321
Qy 1062 CVQJDOQDVEVDEAACAALVRPEASVPC-----LIADC-TYRHVHGTWMEC 1107
Db 1322 CTRGSEG--TIVDFYFCDRNTRPKTKTCEXDTCDGPRVLQKQADVPPIRWATGPWTAC 1379
Qy 1108 SVSCGDIQRRDTCGLGPOAQPAPDFCQHLPKPVTYVGCWAGPC 1153
Db 1380 SATCGNGTQRLLKAC--RDHVRDLDPDEYCNHLDKEVSTENCLRDC 1423
RESULT 4

US-09-369-364A-17
; Sequence 17, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirokawa, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-10
US-09-369-364A-17

Query Match 12.2%; Score 945; DB 4; Length 1081;
Best Local Similarity 24.1%; Pred. No. 6.9e-68;
Matches 306; Conservative 129; Mismatches 433; Indels 400; Gaps 45;

QY 1 PORPTRP-KAPSHAPFLGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWGSHP--- 56
Db 34 PSRHLLPGAAPRHGG-----HSRVPLLQSLASTHFLNLTSSRLLAG 78
QY 57 -----QOSCLQA--LEQAVSSYL----- 73
Db 79 RVSVEYWTREGLAWQRAARPHCLYAGHLQOQASSHVAISTCGGLHGLIIVADEBYLIEP 138
QY 74 -----SP-----GAPLKGRP----- 83
Db 139 LHCGPKGSRSPESGPHCVYKRSRLRPHLDTAGVRDEKFWKGPFWLRLTKPPAPRL 198
QY 84 -----PSPQFORQORQRAAGGILHLELVAVGPDVFOAH-QEDTERYVLTNINIGAE 136
Db 199 GNETERGQGLKRSVSER-----YVEITMDVADKMVAVHGRDVEQYLAIMNIVAK 251
QY 137 LURDPSLGAQFRVHLVKMVLTEPEGAPNITANITSSLLSVCGHSQTI-----NPEDD 189
Db 252 LFQDSLSGSTVNILVTLITLTDQPTLEITHHAGKSLDSFCWKQKSIWNHSGHNAIPE 311
QY 190 TDFGHADLVLIITPDL-----ELPDGNRQVRGVTOLGACSPWTS-CLITEPTDGLGVT 244
Db 312 NGVANHDVAVLITRYDICIYKPKCGT-----LGLA---RWAECSVARSAAASMRTL 360
QY 245 IA-----HEIGHSGFLEHDCGAGCGSGPSGH-----VMASDGAAPRAGLAWSPCRRQLL 294
Db 361 AATSVHCHHEIGHTFGMNDGV-GNSCGARQDPKALNAAHITMKTNPVWSSCNRDYIT 419
QY 295 SLISAGRARCVWDPRPQPGSAGHPDDAQPLIYSANQCRVAFGKXAVACTFAREHLD 354
Db 420 SFLDGSLGLCLANRP---PRQDFVYPTVAPQAYDADEQCRFQHGKSRQCKYK---EV 472
QY 355 QCALSCHTDPLDQSS-CRSLVPLLDGTEC-----GYEK-WCSGRCRSLVELTPIAA--- 405
Db 473 CSELWC-----LSKRNRCITNSIPAEEGLTCQTHIDKGCYKRVK-----VPFGRSREG 522
QY 406 VHGRWSSWGPRSPCSRSCGGVWTRERRCNPRPAPGGRACVAGDLOAEMCNMTQACEKTQ 465
Db 523 VDGAWGPMTPMGDCSRCTCGGVSSSRHCDSPRTIGGKYCLGERRRHRSCTDDCPGS 582
QY 466 LEFMSQCCARTDGOPLRSFGASFYHWAAGVPHSQGDALCRHMCRAIGESFIMKRGDSF 525
Db 583 QDFREVQCAEFDSPFR-----GKSYKM--KTYRGGVKACSLTSLAEGFNFTYERAAV 635
QY 526 LDGTROMPSGPREDGTLSELVSGSCRTFCGDRMDSQQWDRCCQVCGGNSTCSPKRGSP 585
Db 636 VDGTPCRPD-----TVDICSGECKVCHGCDRLGSLREDKCRVCGGDSACETIEGVF 689
QY 586 T-AGRARBYVTLTVPNLTSTVIANHRLPFLTHLAVRIGRYVYVAGKMSISPNITVPSLL 644

US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: 057092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 12.1%; Score 937.5; DB 4; Length 727;
Best Local Similarity 30.1%; Pred. No. 1.6e-67;
Matches 238; Conservative 102; Mismatches 334; Indels 117; Gaps 27;

QY 83 PPSPGFQORQORRAAGGILHLELVAVGPDVFOAHQEDTERYVLTNINIGALLRDP 142
Db 2 PSFGG-----SIRKKRFVSSPRVETMLVADQSMADFHGSLGKXLYLITLFSVAARFYKHP 57

143 LCAQFRVHLVAVKVLTEPAGAPNITANLTSSLLSVCGWSQTINPEDDTPDGHADLVIT 202
58 INNS:SLVVVILVYEEQKPEVTSNAUTLRNFCWQKHNPSDRDPHYDTAULT 117
203 RFDELPLDGNRQVRGVTQLGAGSPTWSCILITDGTDLGVTIAHEIGHSGFLEHGDGAPG 262
118 RQDL-CGSHCTDLGADVGTCDPSCSVIEDDGLQAFTHAELGHVFNMFHDDAKH 176
263 -----SGCGPSGHVWAS-----DGAAPRAGLAWSPCSRQLLSLSAGRARCVMDPPRPOP 313
177 CASLNGVSGDHLWASMLSLDSHQ-----WSPCSAYMVTSLDNGHGECLMDKQPQ--- 228
314 GSAGHP---PDAOPGLYYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSC 370
229 -----NPIKLPDLPGTLTYDANROCFYFESKCPDA-----ASTCTLWCTGTSGGLVVC 281
371 SLLVPLLDGTEGVEKWCGRCSRSLVELTPIAA-VHGRWSSWGPSPCSRSCGGVVT 429
282 QTKHFWADGTCGEGKVCUSGKCNKTKMFKPATPVHSGWPGPWGDCSRTCGGVQY 341
430 RRRQCNPRPAFGACVADGLQAMCNOAC-EKTLQEFMSQOCARTDGOPLRSSPGA 488
342 TWRECDNPPVKNKGKCEGRVYRSCNIEDCPDNGKTFREEQC-EAHNEFSKASFGNE 400
489 SFYHGAAPHSQGDALCRHVCRAIGSFIMKRGDSFLDTRCMPSPREDGTLSLCVSG 548
401 PTVEWTPKYGAVSPKDRCKLTCEAKGIGYFFVLPQKVVGDGTPCSP-----DST-SVCVQG 454
549 SCETPCDGMDSQVWDRQVCGGDNSTCSPRKGSPTAGAREVYVTLTVPNLTSYVI 608
455 QCVYACDRIIDSKKPKDCKGCGGNGSTCKMKGIVTSTRP-GYHDIVTIPAGATNIEV 513
609 ANH-----RELFTHLAVRIG-GRVYVAGKWSISNTTVPVPSLLEDGRVEYRVALTEDRLPR 662
514 KHRNQSGSRNNGSFLAARDGTVLNG-----NFTLSTLEQDLYKGTVLYRSGSAA 567
663 LEEIRWGPLOEDADIQVYRYRYEYGNLTDPDITFTYFQPKRQA-----WV--- 710
568 LEIRSFSPKLEPTIQVLM-----VGHALRPKIKFTYFMKKTSFNAPTFSEWVIEE 622
711 WAAVRGCVSVSCAGLRWVNYCLDQARKELVETVQCO-----GSQOPPAWPEAC 760
623 W-----GECSTCGSG-W-----VQCRDINGHPASECAKEVKPASTRPC 664
761 VLEPCPYNAVDFGPCSASCGGLRERPVYRVEAQSLLKTLPPARCAGAQCPAVAL 820
665 ADLPC-PHWQGVDSWSPCSTCGKGYKRTLXCVSHDGGVLSN-----E 706
821 TCNPQPCPARW 831
707 SCDPLKXPKHY 717

RESULT 6

US-09-491-522-5
Sequence 5, Application US/09491522
Patent No. 6428998
GENERAL INFORMATION:
APPLICANT: Colige, Alain
APPLICANT: Lapierre, Charles M.
APPLICANT: Prockop, Darwin J.
TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
AND THE PRODUCTION, METHODS AND USES THEREOF
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds, LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: USA
ZIP: 10036-2811
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/491,522
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/886,333
APPLICATION NUMBER: 08/886,333
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Abrams, Samuel B
REGISTRATION NUMBER: 30,605
REFERENCE/DOCKET NUMBER: 8389-0060-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-493-4935
TELEFAX: 650-493-5556
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1211 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-491-522-5
Query Match 11.9%; Score 920; DB 4; Length 1211;
Best Local Similarity 27.8%; Pred. No. 8.8e-66;
Matches 283; Conservative 118; Mismatches 370; Indels 248; Gaps 42;
QY 91 QROQRRAAGGILHLELLVAVGPPVFOAH-QEDTERVYVLTNLCAGELLRDSLSGQFRV 149
DB 254 RRRARRHAADDYNEIIVLLGVDDSVQFHGKEHVQKYLTLNMIYVNIYHDSLSGAHNV 313
QY 150 HLYKMWILTEPEGAPNI-TANLTSSLLSVCGWSQTINPEDDTPDGHADLVITREDEL 208
DB 314 VLVRITLLSYKGSVSLIEIGNPSQSLNVCKWAVLQKPDGTGHDYHDAIHLTRQDF-- 371
QY 209 PDGNROVRGVTQCGGACPTWSCILITDGTDLGVTIAHEIGHSGFLEHGDGAPSGCGPS 268
DB 372 --GPSGMOGYAPVTGMCHPVRSCITLNHEDGSSAFVVAHEIGHVGLGMEHDG-QGNECGDE 428
QY 269 ---GHVNASDGAAPRAGLAWSPCSRRLLSLRSAGRARCVMD-----PPRQPGSAGH 318
DB 429 VRLGSIMAPLVQAAFRHFHWSRCSQBELSYLHS--YDCLLDDPFAHDWPAFPQ----- 480
QY 319 PPDAPQGLYYSANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDPLDQSSCSRLLVPL 378
DB 481 ----LPLGLHYSNNEQCRFDGLGYMMCTAFET-FDPCQLWC-SHPDNPYFCRTKKGPP 534
QY 379 DGTCEGVKWCGRCSRSLVELTP-IAAVHGRWSSWGPSPCSRSCGGGVVTRRRQCNPN 437
DB 535 DGTWCAPGKCFKHC---IWLTPDILKRDGSGWGSWSPGSCSRCTGTVKTRQCDNP 591
QY 438 RPAFGGRACVADLQAEWCNTQACEKTQLEFMSQOCARTDGOPLRSSPGGAFYHGAAY 497
DB 592 HPANGGRTCSGLAYDFOLCSRQDCPDLSLADFREEQCROWD---LYFEHGDQ-HHW---L 644
QY 498 PHSQGD--LCRHW--RAIGESFIMKRGDSFLDGTCTCMPSPREDGTLSCVSGSCRTF 553
DB 645 PHEHRAKERCHLYCESRETGEVVSMMKR--MVHDTGTRC-----SYKDAFSLOWRGDCRK 697
QY 554 GCDGRMDSQWDBRCQVCGGDNSTCSPKSGFT-AGRAREVYVTLTVPNLTSVYIANHR 612
DB 698 GCDGVIGSKQEDKCGVCGGDNSHCKVVKVGTFTSPKXHYKMEFIPAGARHLLIQEVD 757
QY 613 PLFTHLAVR--IGGRYVY-----AGKMGISPNNTYPSLLEDGRVEYRVALTEDLPR 662
DB 758 ATSHLAVKMLETKFILNEENDVDASKTIAMGVWEYREDGE-----803
QY 663 LEEIRWGPLOEDADIQVYRYRYEYGNLTDPDITFTYF-----OPKPRQAW 709

Db 804 -BTLOTMGLHGTITVLVI-----PVGD-TRVSLTYKMIHEDSLNVDDNNVLBEDSVVY 856
Qy 710 VMAAVR-GPCSVSCGAGLRWVNYSCLDQARKELVETVQCQSQSQPPAWPEACVLEPC-PP 767
Db 857 EWALKWSPCKSCGCGSQFTKYGCRRLDHKMVRHGFCAALSKPKAIRACNPDCECQP 916
Qy 768 YMAVDFGPCSACG-GGLRERPVRCVE-AQGSLLKTLPPARCAGAGQPAVALETCPQ 825
Db 917 VMTGWEPSCQTCGRGTGMQVRSVRCIQPLHNDNTRSHAKHCNDARPE---SRRACSR 973
Qy 826 PCPARVESEPSCTSGAGLALENETCPVGDGLEAPVTEGSGVDEKLPAPPCVGM 885
Db 974 LCPGRWA----- 981
Qy 886 SCPPGWHLDGAGEKAPGWSGIRTAQAHHVMTPAAGSCSVSCGRGLMELRPLC--M 943
Db 982 -----GPM-----SQCSTCGNGTQERVLCTA 1005
Qy 944 DSALRVVPEELCGLASKPSRRREVCOAVPCPARWQYKLAACSVSCGRVRRILYCRA 1003
Db 1006 DDSFGI-CQEB-----RPEAR-TCRLGPCP-----RNI----- 1032
Qy 1004 HGEDDEEILLTQCQLPRPEQACSLPCPRWKVMSLPGCSASCGLGTARSVAC 1062
Db 1033 --SDPSKXSV--VQWLSRPDPDS-----FIRKISSKHCOGD-----KSIFC 1071

RESULT 7

US-09-491-522-11

; Sequence 11, Application US/09491522

; Patent No. 6428998

; GENERAL INFORMATION:

; APPLICANT: Colige, Alain

; APPLICANT: Lapiere, Charles M.

; APPLICANT: Prockop, Darwin J.

; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,

; AND THE PRODUCTION, METHODS AND USES THEREOF

; NUMBER OF SEQUENCES: 17

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds, LLP

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036-2811

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows

; SOFTWARE: FastSeq for Windows Version 2.0b

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/491,522

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/886,333

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Abrams, Samuel B

; REGISTRATION NUMBER: 30,605

; REFERENCE/DOCKET NUMBER: 8389-0060-999

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-493-4935

; TELEFAX: 650-493-5556

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1205 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-491-522-11

Query Match

Best Local Similarity 11.9%; Score 918.5; DB 4; Length 1205;

Matches 254; Conservative 105; Mismatches 339; Indels 101; Gaps 32;
Qy 91 QROQRRAAGGILHLELLVAVGPDVFOAH-QEDTERYVLTNLNICAELLRDPSSLGAQPRV 149
Db 248 RMRMRHAADDYNNIEVLGVDSDVQFHGTEHVQKYLILMNIWNEIYHDESGLAHNV 307
Qy 150 HLYKMWILTEPGAPNI-TANLTSSLLSVGWSQ-TINPEDDTPGHADLVLYITRFDE 207
Db 308 VLVRILLSVGKMSLIEIGNQSLENVCRWAYLQOKPDTDHDYH-DHAIFLTRQDP- 365
Qy 208 LPDGNQVRGVTOLGGACSTNSCLITEDTGFDLGVITIAHEIGHSGFLHGDGAPGCGP 267
Db 366 ---GPSGMOQYAPVTGMCHPVRSCITLNHEDGFSAFVVAHETGHVLMGHDG-QGNRCDD 421
Qy 268 S---GHVMSADGAAPRAGLAWSPCSRRLQLLSLSAGRARCWMDPPRPOFGSAGHPDPAQ- 323
Db 422 EYVLGIMAPLVQAARHFRHWSRCSQELRYLHS-----YDCLRDDPFFHDMALPQL 475
Qy 324 PGLYISANEOCRVAFGKAVACTFAREHLDWCQALSCHTDPLDSSSRLLVPLLDGTBC 383
Db 476 PGLHYSMNEOCRDFGLGYVMCTAFRT-FDPCKQLWC-SHPDNPYFCKTKKGFLDGTMC 533
Qy 384 GVBKWSKGRCSRSLVELTPIAAVHGRWSSWCPSPCSRSCGGVVTRRCCNNPRPAFG 443
Db 534 APGKHCFKGHCTWLT--FDILKRDGNWGWSPFSCSCTCGTGKFRTRQCNPHPPANG 591
Qy 444 RACVGADLQAEWCNTQACEKTOLEFMSQCCARTDQQLPSRSPGGGASFYHMGAAVPHSQD 503
Db 592 RTCSGLAYDFOLCNSQDCPALADFREECQKQWD---LYFEGDQD-HHW---LPREHRD 644
Qy 504 A--LCRHMV--RAIGESFIMKRGDSFLDGTGRCMPSPGREDGTLSLCVSGSCTFGCDGRM 559
Db 645 AKERCHLYCESKETGEVVSVMKR--MVHDGTRC-----SYKDAFSLCVRGDCRKGVCQDGI 697
Qy 560 DSQVWDRQVCGGDNSTCSPKGSFTAG-RAREYVTFVTVPNTTSVIANHRPLFTHL 618
Db 698 GSSKQEDKCGVCGGDNHCKVKGTFGRSKFKLYIKMFEIPAGARHLLIOEADTTSHL 757
Qy 619 AVR--TGRVTVAGKMSISNT-TYPSL-----LEDGRVEYRYVALTEDRLPRLEIRI 668
Db 758 AVKNLETGKFIENEDVDNPKSTFIAMGVWEVYRDEDR-----ETLQT 802
Qy 669 WGPL-----QEDADIQVRRY-----GEEYGNLTRPDITFTYFQPKPQAVW 711
Db 803 MGPLHGTTTVLVIPEGDARISLTVMYMHEDSLNVDDNNVLEDDSV-----GYEW 852
Qy 712 AAVR-GPCSVSCGAGLRWVNYSCLDQARKELVETVQCQSQSQPPAWPEACVLEPC-PPY 769
Db 853 ALKKWSPCKSCGCGSQFTKYGCRRLDHKMVRHGFCDVSXKPKAIRRTCNPOECQPVW 912
Qy 770 AVGDFGPCSACG-GGLRERPVRCVE-AQGSLLKTLPPARCAGAGQPAVALETCPQPC 827
Db 913 VTGEWEPSCSRGRTGMQVRSVRCVQPLHNDNTRSHAKHCNDARPE---GRRACNREL 969
Qy 828 PARVESEPSCTSGAGGAG 846
Db 970 PGRWAGSWSCSVTCGNG 988

RESULT 8

US-09-445-023A-1

; Sequence 1, Application US/09445023A

; Patent No. 6565858

; GENERAL INFORMATION:

; APPLICANT: Hirose, Kunitaka

; APPLICANT: Inoguchi, Eiji

; APPLICANT: Hakozaaki, Michinori

; APPLICANT: Ishioaka, Keiko

; APPLICANT: Ishida, Yukako

; APPLICANT: Matsushima, Kouji

; APPLICANT: Kuno, Kouji

; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical		; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical	
; FILE REFERENCE: Q57092		; FILE REFERENCE: Q57092	
; CURRENT FILING DATE: 1999-12-03		; CURRENT FILING DATE: 1999-12-03	
; PRIOR APPLICATION NUMBER: JP 9-160422		; PRIOR APPLICATION NUMBER: JP 9-160422	
; FILING DATE: 1997-06-03		; FILING DATE: 1997-06-03	
; NUMBER OF SEQ ID NOS: 14		; NUMBER OF SEQ ID NOS: 14	
; SOFTWARE: PatentIn version 3.0		; SOFTWARE: PatentIn version 3.0	
; SEQ ID NO 1		; SEQ ID NO 1	
; LENGTH: 727		; LENGTH: 727	
; TYPE: PRT		; TYPE: PRT	
; ORGANISM: Homo sapiens		; ORGANISM: Homo sapiens	
US-09-445-023A-1		US-09-445-023A-1	
Query Match		Query Match	
Best Local Similarity 29.9%; Pred. No. 1.2e-65;		Best Local Similarity 29.9%; Pred. No. 2e-65;	
Matches 236; Conservative 111; Mismatches 314; Indels 127; Gaps 30;		Matches 238; Conservative 111; Mismatches 319; Indels 134; Gaps 31;	
QY	90 RORQRRAAGGILHLELLVAVGPDVFOAQHEDTERVYLTNIGAEHLRDPISLGAQPRV 149	QY	83 PESPQFORQ-----RORRAAGGILHLELLVAVGPDVFOAQHEDTERVYLTNIGCA 135
DB	5 RIRLRKRVSPPRYVETMLVADQSMAPFSGGLKHYLLTLFSAARLYKHPSTRNSVSL 64	DB	231 PQPALQGVGPTGTGTSIRKKRVSSHRYVETMLVADQSMAPFSGGLKHYLLTLFSAVA 290
QY	150 HLKVMILTEPEGAPNITANLTSLLSVCGWSQINPEDDTPGHADLVLTITFDLELP 209	QY	136 ELIRDPSLGAQFRVHLVWVILTEPEGAPNITANLTSLLSVCGWSQINPEDDTPGHA 195
DB	65 VVKILVIHDEQKPEVTSNAALTRNFCWQKQHNPPSDRDAEHYDTAILFTRODL--- 121	DB	291 RLYKHSIRNSVSLVVKILVIHDEQKPEVTSNAALTRNFCWQKQHNPPSDRDAEHY 350
QY	210 DGNR--QVRGVTOLGACSTWCLITETGFDLGVTTIAHIGHSFGLHDPAGP-----S 263	QY	196 DLVLYITRFDLELPDGNR--QVRGVTOLGACSTWCLITETGFDLGVTTIAHIGHSF 253
DB	122 CGSQTCDTLGMADVGTCDPSRSCSVIEDGLQAFTAHELGHVFNPHDADKQACASLN 181	DB	351 DTAIFTRQDL---CGSQTCDTLGMADVGTCDPSRSCSVIEDGLQAFTAHELGHVF 407
QY	264 GCGPSSHVMAS-----DGAAPRAGLAWSPCRRQLLSLSAGRACVWDPRPQPGSAGH 318	QY	254 GLEHDPAGP-----SGCPSCHVMAS-----DGAAPRAGLAWSPCRRQLLSLSAGRAC 304
DB	182 GYNQDSHMMASMLSNLDHSP-----WSPCSAYMITSFLDNGHECLMDKQP-----N 229	DB	408 NMPHDDAKQACASLNGVNDHMMASMLSNLDHSP-----WSPCSAYMITSFLDNGHEC 462
QY	319 P---PDAQGLYISANEQCRVAGPKAVACTFAREHLDMAALSCHTDPLQSSCSRLLV 375	QY	305 VMDPRPQPSAGHP---PDAQGLYISANEQCRVAGPKAVACTFAREHLDMAALSCH 361
DB	230 PQLPGDLPTLYDANRCQFTFGEDSKHCPDA---ASTCSTLWCTGSGVLVQTKHF 286	DB	463 LMDKQP-----NPIQLPGDLPTLYDANRCQFTFGEDSKHCPDA---ASTCSTLWCT 512
QY	376 PLLDGTGECGVEKWCGRCSLVEL-----TPIAAVHGRWSWGSPRSCSGGVWTRR 431	QY	362 TPLDQSSCSRLLVPLLDGTGECGVEKWCGRCSLVEL-----TPIAAVHGRWSWGSPRS 417
DB	287 PWADGTSCEGKWCINGKCNKTDKHFDTFF---HSGWGPWPGWDCSRICCGGVQYTM 343	DB	513 GTSGGVLVCTKHFPWADGTSCEGKWCINGKCNKTDKHFDTFF---HSGWGPWPGW 569
QY	432 RQCNPRPAPFGGRACVAGDIQAEMCNTOAC--EKTQLEFMSQOCARTDQPLRSSPGGASF 490	QY	418 PCSRSCGGVYVTRRRQCNPRPAPFGGRACVAGDIQAEMCNTOAC--EKTQLEFMSQOCART 476
DB	344 RECDNVPKNGKYCEKRVYRSCNLEDCPDNNGKTFRESQC--BAHNEFSKASFGSPA 402	DB	570 DCSRTCGGVQYVTRRECDNVPKNGKYCEKRVYRSCNLEDCPDNNGKTFRESQC--EA 628
QY	491 YHWGAAPVHSGDA---LCHRMCRAGISFIMKRGDSFLDGTGRCMPSPREDGTLSLVS 547	QY	477 DGOPLRSSPGGAFYHWGAAPVHSGDA---LCHRMCRAGISFIMKRGDSFLDGTGRCMP 533
DB	403 VEW---IPKYGAVSPKDRCKLICQAKGIGYFFVLQPKVVDGTPCSP-----DST-SVCVQ 453	DB	629 HNEFSKASFGSGPAVEW---IPKYGAVSPKDRCKLICQAKGIGYFFVLQPKVVDGTPCSP 685
QY	548 GSCRTGCDGMDSQVWDRQVCGGDNSTCSPRKGSFTAGREYVTLVTPNLTSTY 607	QY	534 SGPREDTLSICVSGSCRTGCDGMDSQVWDRQVCGGDNSTCSPRKGSFTAGREY 593
DB	454 GQCVKAGCDRIIDSKKFKDKGCVGCGNGSTCKKISGSVTSAPK-GYHDIVTIPTGATNIE 512	DB	686 -----DST-SVCVQGVKAGCDRIIDSKKFKDKGCVGCGNGSTCKKISGSVTSAPK-GY 738
QY	608 IANH-----RPLFTHLAVRIG-GRYVWAGKMSISNTTYPSSLLEGRVEYRVALTEDRLP 661	QY	594 VTFLTVPNLTSTVVIANH-----RPLFTHLAVRIG-GRYVWAGKMSISNTTYPSSLLEGR 647
DB	513 VKORNGSRNNGSFLAIKAADGYILNGDYTL-----TLEQIMYGVVLRVSGSSA 566	DB	739 HDITITGATNIEVKORNGSRNNGSFLAIKAADGYILNGDYTL-----TLEQDI 792
QY	662 RLEIRIWIPLQEDADIQVYRYGEEYGNLTPDITFTYFQPKRP-----AWV--- 710	QY	648 RVEYRVALTEDRLPRLEIRIWIPLQEDADIQVYRYGEEYGNLTPDITFTYFQPKRP 707
DB	567 ALERISFSLKEPLTIQVL-----TVGNALRPKIKYTFVFKKXESFNAITFSAWVIE 621	DB	793 MYKGVVLRVSGSSAALERISFSLKEPLTIQVL-----TVGNALRPKIKYTFVFKKXKE 847
QY	711 -WAAVRGPCSVSCAGLRWNVNYSCLDQARKBELVTVQCGSQOP-----PAWPEAC 760		
DB	622 EW-----GECSSCELG-W-----QRRLVECRDING--QPASECAKEVKPASTRPC 664		
QY	761 VLEPCPYWAGDGPSCASCGGLRERPRVCEAQGSLLKTLPPARCAGAAQPPAVAL 820		
DB	665 ADHPC-FQWQLGEWSSCKTCGKGYKRSKLCLSHDGGVL-----SHE 706		
QY	821 TCNPQPCP 828		

QY 708 -----AWV---WAAVRGPCSVSCGAGLRWVNYSCLDQARKELVETVQCQSQOP-- 753
Db 848 SFNAIPTPSAWVIEWE-----GECSSKCELG--W-----QRLVECRDING--QPAS 890
QY 754 -----PAMPEACVLEPCPPYMAVGDFGFCPSASCGGLRERPVRCVVEAQSLLKTLPPA 806
Db 891 ECAKEVKPASTRPCADHPC-PQWLGEWSSCSKTGKGYKRSJKLSHDGGVL----- 943
QY 807 RCRAGAQQPAVALETCPQPCP 828
Db 944 -----SHESCDPLKCP 954

RESULT 10
US-09-568-559-2
; Sequence 2, Application US/09568559
; Patent No. 6649377
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Allard, John
; APPLICANT: Heller, Renu
; APPLICANT: Van Wart, Harold
; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/09/568,559
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-09-568-559-2

Query Match 11.8%; Score 912; DB 4; Length 949;
Best Local Similarity 29.7%; Pred. No. 2.8e-65;
Matches 238; Conservative 110; Mismatches 320; Indels 134; Gaps 31;
QY 83 PPSPGFQRORQ-----RORRAGGTLHLLELVAVGPDVFAHQEDTERVYLTNLNIGA 135
Db 213 PQDPALQGVQPTGTGIRKRFVSSHRYVETMLVADQSMAEFGSLKHYLLTLFVAA 272
QY 136 ELLRDPSLGAQFRVHLVXVILTEPEGAPNTANLTSSLVSCWGSQINPEDDTPGHA 195
Db 273 RLYKHPISIRNSVSLVVKILVIHDEQKGPVTSNAALTLRNFNCWQKQHNPFSDRAEHY 332
QY 196 DIVLYITRFDLELPDGNR--QVRGVTLGGACSPWSCCLITDGTDLGVTAHGHSHF 253
Db 333 DTAILFTRQDL---CGSQTCDTLGMADVGTVCDFSRSCSVIEDDGLQAAFTTAHELGHVF 389
QY 254 GLEHHDGAPG---SGCPSGHWAS-----DGAAPRAGLAWSPCRRQLLSLSAGRARC 304
Db 390 NMFHDDAKQACSLNGVNDQSHWASMLNLHDSQP-----NSPCAYMITSLDNGHGC 444
QY 305 WDPFPPRQPGSAGHP---PDAQPGLYSANGQCRVAFOPKAVACTAREHLDMCOALSCH 361
Db 445 LMDKQP-----NPIQLPGLPGTSDANRCQCTFGEDSKHCPDA---ASTCSTLWCT 494
QY 362 TDFLDQSCSRLLPLLDGTGCGVEKWCCKGRCSLVEL-----TPIAAVHGWSWNGPRS 417
Db 495 GTSGGVVLVQTKHPFADGTSCGEGKWCINGKCNKTRKHFDTPF---HGSWMGMGPGW 551
QY 418 PCSRSCGGVYTRQCNRPAPFGACVAGADLAQEMCNTQAC-EKTQLEFMSQOCART 476
Db 552 DCSSTCGGVGYTRWECNDPNVPKNGKYCEGRKRVYRSCNLEDCPDNNKGTTFREQC-EA 610
QY 477 DGOPLRSPPGASFYHGAAPVHSGQDA---LCRHMCRAGIBSFIMKRGDSFLDGTTCMP 533
Db 611 HNEFSKASFGSGPAVEW---IPKYAGVSPKDRCKLICQAKGIGYFVLPQKVVDTGTPCSP 667

QY 534 SGPREDGTLSLVSGSCRTFGCDGRMDSQVMDRQVCOVCGDNDSTCPRKGSFTAGRAREY 593
Db 668 -----DST-SVCVQOCQVAGCDRIIDSKKFDKCGVCGNGSTCKKISGSVTSAPK-GY 720
QY 594 VTFLVTPLNTSVYIANH-----RPLFTHLAVRIG-GRYVYVAGKMSISPNTPYPSLLEDG 647
Db 721 HDIITPITGATNIEVKQRNQRNRNGSFLAIKAADGTVILNGDYTLN-----TLEQDI 774
QY 648 RVEYRVALTEDRLPLEBIRINGPLOEDADIQVRRYRGEYGNLTRPDITFTYFQPKPRQ 707
Db 775 MYKGVVLRYSGSSAALERIRSFPLKEPLTIQVL-----TVGNALRPKIKYTYFVKKKKE 829
QY 708 -----AWV---WAAVRGPCSVSCGAGLRWVNYSCLDQARKELVETVQCQSQOP-- 753
Db 830 SFNAIPTPSAWVIEWE-----GECSSKCELG--W-----QRLVECRDING--QPAS 872
QY 754 -----PAMPEACVLEPCPPYMAVGDFGFCPSASCGGLRERPVRCVVEAQSLLKTLPPA 806
Db 873 ECAKEVKPASTRPCADHPC-PQWLGEWSSCSKTGKGYKRSJKLSHDGGVL----- 925
QY 807 RCRAGAQQPAVALETCPQPCP 828
Db 926 -----SHESCDPLKCP 936

RESULT 11
US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurekainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-7
US-09-369-364A-7
Query Match 11.2%; Score 864.5; DB 4; Length 997;
Best Local Similarity 29.3%; Pred. No. 2.2e-61;
Matches 275; Conservative 94; Mismatches 380; Indels 191; Gaps 36;
QY 17 LGLALLRMHQRHPRARCPPLC-----VAGILACGFLLCGWPSPSHFOQSLCQ 62
Db 116 LGRANRAH-----TPACHLLGEVQDPELEGALAIASACDGLKGVFLSN-EDYFIE 166
QY 63 ALR-POAVSSYLSFGAPLKGRRPPSPGFOR-----QEQOR 96
Db 167 PLDSAPARPGHAQPHVYVYKQAPERLAQRGSSAPSTCGVYVYPELESRRERWEQQRW 226
QY 97 RAAGGTLH-----LELLVAVGPDVFOAH-QEDTERVYLTNINIGALLRDPSSLGAQ 146
Db 227 RPLRLRHQRSVSKKWCETLVADAKMVEYHGOQVESYVLTIMMWAGLFHDPISGNP 286
QY 147 FRHLVKMWILTEPEGAPNTANLTSSLVSCWGSQINPEDDTPGHADLVLYITRFDL 206
Db 287 IHITVRLVLEDEEDLKITHHADNTLKSCKWCKSNMKGDAPLHHDHTAILLTRKOL 346
QY 207 ELPDGNR--QVRGVTLGGACSPWSCCLITDGTDLGVTAHGHSHFGLHHDGAPGSG 264
Db 347 -CAAMNRPCELTGLSHVAGMCQPHRSCINEDTGLPLAFTVAHELGHSGFIQHDGS-GND 404
QY 265 CGPSG---HVNASGGAAPRAGLAWSPCRRQLLSLSAGRARCVDWPPRPPQSGSAGHPD 321
Db 405 CEPVGKPFIMSPQLLYDAAPLTWSRCSQVITREFLDKRWGLCLDDPPAKDIIDF---PS 461

322	QY	ADPGLYSANEQCRVAFGPKXAVACTTFAREHLDMCOALSCHTDPLDSCSSRLVPLPLDG	380
462	Db	VFPGLVYDVSQCFLQYGAISAFCEMDNVCHTLWCSVGTTTCHSKLDAADV	513
381	QY	TTGCVKXKSGKGRCSLVELTPTA---AVHGRSSWGSPRSPCSRSCGGVVTERRQCN	436
514	Db	TRCGENKWLISGEC---VPVGRPEAVDGGWSGSAWICSRSRSGMGVQSAERQCTQ	567
437	QY	BPAPFGRACVGADLQAEMCNTQACEKTOLEFNSQQCARTDGQPLRSSPGGASVYHGA	496
568	Db	PTPKYKRGYCVGBERKPRFLCNLQCAPAGRSFRHVQCSHFDAMLYK-----GQJLHW---	619
497	QY	VPHSQDALCRHMCRAIGESFIMKRGDSDLGTCTMPSGPPREDGTLSLVSGSCRTFGCD	556
620	Db	VPVNDVNCLEHCRPANEFPAKLRDACVDGTFCYQVRSRDL---LCINGICKNVGCD	675
557	QY	GRMDSQWDRQVCGGDNSTCSPRKSGFTAGRAREYVTFVTVPNTLTSVTYIANHRPLFT	616
676	Db	FEIDSGAMEDRCVCHGNGSTCHTVSGTFEEAEGLYVDVGLIPAGAREIRIQEVAEAA	735
617	QY	HLAVRTGG--RYVVAGKMSIPNTTYSLLDEGRVEYRVA---LTDRLPLREHIRINWP	671
736	Db	FLALRSDEPEKTYFLNGGWTIQWNG-----DIQVAGTTTTFARRGNWENLTSFPG	784
672	QY	LQEDADIQVYRVE---EYGNLTRPDIITYFOP---KPEQAWMAAVRGP---CSVSC	722
785	Db	TRKPVNIQVPASRGPGGSRGGVPRPSTLHGRSRPGVSPGSGVTEPGSEPGPPAAASTV	844
723	QY	GAGLRWNYSCLDQARKELVETVQCQSGQPP-----AWPEACVL-----	762
845	Db	SFSLKWFN-----LVAHVHGGWGQAPLGLGGRHRLVLMGPLPTQLLFOESNP	894
763	QY	-----EPCPPY--WAYVDGFGPCSASCGGL-----RERPVRCE---VEAQGS	798
895	Db	GVHYEYTIHREAGGHDEVPVPVFSVHWYGPWKCTVTCGRGEKWRHSPT-CRGLVSGQGH	953
799	QY	LLKTLPPARCAGAQQPAVALETNCPOP-----CPARNEVS	834
954	Db	WLQL--PAHCWA-----TTGLEVCFSPEQFSICEMRLAIA	986

```

RESULT 12
US-09-800-729-89
; Sequence 89, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044Pl
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 1745
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-89

Query Match 11.1%; Score 857.5; DB 4; Length 1745;
Best Local Similarity 24.6%; Pred. No. 1.8e-60;
Matches 275; Conservative 152; Mismatches 434; Indels 257; Gaps 41

QY 408 GRVSSWGPBPCCSRGGVVTTRQCNPRPAPGGRACVAGADLAQEMCNTOACEKTOLE 467
DB 34 GLMDAWGEWSECSRTCGGASYSLR-----LSKSCGEGNIYRTCSNVDCPPEAG 87
OY 468 FMSOCCA-----RTDGPRLRSPGGASFYHWGAAPVHSGO-DALCRHMCRAIGESFIMKR 521

```

[illegible]

RESULT 13

[illegible]

593 HIEVREVMKSNVIALKSEGGDYVINGAWTID----WPRKFDVAGTAFHYKRTDPEPSL 648
664 EEIRINGPLOEDADIQV---RRYGEYV---NLTRP-----DITFTYFQKPRQAWVA 712
649 EAL---GPTSENLLVMVLLQEQNLGIRYKFNVPITRTSGDNEVGFTW---NHQPW--- 698
713 AVRGPCSVSCGAG-----LRWVNYSCLDQA---RKELVETVQCQ 748
699 ---SECSATCAGGKWPTRQPTQARWRRTKHILSYALCLLKLLIGNISCR 744

Search completed: March 13, 2004, 07:46:05
Job time : 64.7375 secs

714 VRGPCSVSCAGLRWVNYSCLDQARKELVETVQCQSQCPAPWPCVLEPC 765
559 --GECSTCKGKYKRSKLKSLSHDGVLSHE-SCDPLKPKXHFIDFCTMAEC 607

RESULT 15
US-09-963-791-24
; Sequence 24, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Doncho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399a1 Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-24

Query Match 10.0%; Score 776; DB 4; Length 757;
Best Local Similarity 29.9%; Pred. No. 2.3e-54;
Matches 212; Conservative 106; Mismatches 283; Indels 108; Gaps 32;

94 QRRAAGLHLELLVAVGPVQAH-QEDTERVYLTNINIGAEILLRDPISLGAQFVHLV 152
90 QRKRSVIERFVETLLVADKXMGVGHGRKDIEHYILSVNVIYAKLYRDSLGNVNIIVA 149
153 KWLITPEGAPNITAN-LTSSILSVCGWSOTI---NPEDDTDP-----GHADLVLYITR 203
150 RLIVLTEDQ--PNEINHADKSLDSFKWQKSLSHQSDGNTIPENGIAHHDNAVLIIR 207
204 FDL----ELPDGNRQVRGVTOLGGACSTWISCLITETDGLGVTIAHIGHSFGLHDG 259
208 YDICTYKKNKPG---TLGLASVAGWCEPERSCSINEDIGLSAFTIAHIGHFNFGNHDG 264
260 APGSGCGPSGH-----VMASDGAAPRAGLAWSPCSRQLLSLISAGRARCVWDPPRPQGS 315
265 I-GNSCGTKGHEAAKMAAHITANTNPFSSWSACSRDYITSLDSGRGTCLDNEP---PKR 320
316 AGHPDPAQPGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSS-CSRL 374
321 DFLYPAPVAPGVYDADEQCFQYGATSRCKYK---EVCRELWC---LSKSNRCVTNS 372
375 VPLLDGTGEC---GVBEK-WCKGRCSRSLVELTPIA----AVHGRSSWSGPRSPCSRSCGGG 426
373 IPAAEGTLCTQGNIEKGWYQDCC-----VFGTWPSQIDGGWPSLWGECSRTCGGG 426
427 VVTRRRQCNPRPAPGRACVAGADLAQEMCNTQACEKTQLEFMSQQCAQTGQFLRSSPG 486
427 VSSSLRHCDSPAPSGGKVCIGERKYRSCNTDPCPLGSRDPREKQCADFDNMPFR---- 482
487 GASFYHWGAAVPHSGDAL-CRHMCRAGIESFIMKRGDSFLDTRCMPSGPREDDGLSLC 545
483 -GKYNNW--KPYTGGGVKPCALNCLAGYNFYTERAPAVIDGTQC-----NADSLDIC 532
546 VSGCRTGCGDRMDSQQWDRCOVCGGDNSTCSPRKGSFTAGRAR-EYVFTLTVTENLT 604
533 INGECKHVCGDNILGSDAREDCRVCGDGTDAIEGFNDSLPRGGYMEVVQIPRGSV 592
605 SVYTANHRPLFTHLAVRI-GGRYVWAGKMSISNTTYPSSILEDGRVEYRVALTEDRLPRL 663

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:33:19 ; Search time 3.3849 Seconds
(without alignments)
2756.533 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534
Sequence: 1 PGRPTPKAPSHSAPLLGLA.....PLKGRPPSPGFORQRR 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: Piri:.*
2: Piri:.*
3: Piri:.*
4: Piri:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.5	14.9	353	JC4875	zinc-finger protei
2	74.5	14.0	2337	T40577	probable phosphat
3	74	13.9	220	I80329	receptor tyrosine
4	74	13.9	910	A53137	tyrosine kinase re
5	73.5	13.8	647	AG2114	hypothetical prote
6	72	13.5	402	I46053	connexin44 - bovin
7	72	13.5	1366	B86292	F7H2.12 protein -
8	71.5	13.4	837	A57542	p96 protein - mous
9	71	13.3	913	A48280	receptor tyrosine
10	70.5	13.2	135	E75354	hypothetical prote
11	70.5	13.2	606	T11909	NADH2 dehydrogenas
12	70.5	13.2	771	B38252	granulocyte colony
13	70	13.1	513	JQ1486	activin receptor I
14	70	13.1	513	A39836	activin receptor p
15	70	13.1	513	A49193	type II activin re
16	70	13.1	513	S27258	activin receptor t
17	70	13.1	513	I45850	activin receptor t
18	69	12.9	203	AH2144	hypothetical prote
19	69	12.9	258	B37252	insulin-like grow
20	69	12.9	535	T17212	hypothetical prote
21	69	12.9	920	B34493	collagen alpha 1(I
22	68.5	12.8	1074	T24877	hypothetical prote
23	68.5	12.8	1076	T24877	hypothetical prote
24	68	12.7	300	T24482	hypothetical prote
25	68	12.7	550	I45504	mycolysin (EC 3.4.
26	68	12.7	684	TAS6154	Abi substrate ena
27	68	12.7	2783	I441948	alpha-fetoprotein
28	67.5	12.6	332	T48691	regulatory protein
29	67.5	12.6	499	S09880	hypothetical prote

ALIGNMENTS

RESULT 1

JC4875

zinc-finger protein Zic4 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 05-Nov-1999

C:Accession: JC4875

R:Aruga, J.; Yozu, A.; Hayashizaki, Y.; Okazaki, Y.; Chapman, V.M.; Mikoshiba, K.

Gene 172, 291-294, 1996

A:Title: Identification and characterization of Zic4, a new member of the mouse Zic ge

A:Reference number: JC4875; MUID:96269420; PMID:8682319

A:Accession: JC4875

A:Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-353 <ARU>

A:Cross-references: DDBJ:D78174; NID:gl480004; PIDN:BAAL1240.1; PID:dl011902; PID:gl148

C:Genetics:

A:Gene: Zic4

A:Map position: 9

C:Keywords: zinc finger

F:119-298/Region: zinc fingers

Query Match 14.9%; Score 79.5; DB 2; Length 353;

Best Local Similarity 27.9%; Pred. No. 1.3;

Matches 31; Conservative 12; Mismatches 49; Indels 19; Gaps 4;

QY 1 PGRPTPKAPSHSAPLLGLALL-----RMQBRH-----PRARCPLCVAGILACGF---- 46

DB 51 PGLHEQPPQASHSRPLNGLRLGIPDMYARSEFPAPGPMARSDTLATATALH-GYGMN 109

QY 47 ----LLGCMGSPSHFQOQSCIQALEPQAVSSVLSGAPLKGRPPSPGFORQRQ 93

DB 110 LTWNLTAFHGFGAFFRYMRQPIKQELICKWLGDDSPSPRCSKTFSTWHE 160

RESULT 2

T40577

probable phosphatidylinositol 3-kinase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 15-Sep-2000 #text_change 06-Oct-2000

C:Accession: T39913; T40577

R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Oliver, K.; Harris, D.

submitted to the EMBL Data Library, March 1999

A:Reference number: Z21890

A:Accession: T39913

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 365-2337 <LYN>

A:Cross-references: EMBL:AL049558; PIDN:CAB40167.1; GSPDB:GN00067; SPDB:SPBC216.07C

R:Seeger, K.; Harris, D.; Wood, V.; Rajandream, M.A.; Barrell, B.G.

submitted to the EMBL Data Library, January 1999

A:Reference number: Z21938

A:Accession: T40577

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-398 <SEE>
A;Cross-references: EMBL:AL035216; PIDN:CAA22805.1; GSPDB:GN00067; SPDB:SPBC646.01c
A;Experimental source: strain 972h-; cosmid c646
C;Genetics:
A;Gene: SPBC646.01c
A;Map position: 2
C;Superfamily: yeast TOR2 protein

Query Match 14.0%; Score 74.5; DB 2; Length 2337;
Best Local Similarity 29.9%; Pred. No. 23;
Matches 20; Conservative 10; Mismatches 28; Indels 9; Gaps 3;

QY 17 LGALLRMQHRRPRRCPPCLVA-----GILACGFLGCGP--SHFQOQCLQALEPQA 68
DB 1137 LGVALLRSPSHALRACAAALAAVQPLARLDFNASF-VSCWSELYDFHFEELVKSIEIAL 1195

QY 69 VSYLSP 75
DB 1196 TSPHISP 1202

RESULT 3
I80329
C;Species: Rattus sp. (rat)
C;Status: preliminary; translated from GB/EMBL/DBJ
A;Title: Receptor tyrosine kinase - rat (fragment)
A;Reference number: I59442; MUID:95320273; PMID:7597135
A;Accession: I80329
R;Sakuma, S.; Sava, H.; Ijichi, A.; Tofilon, P.J.
Radiat. Res. 143, 1-7, 1995
A;Title: Radiation induction of the receptor tyrosine kinase gene Ptk-3 in normal rat as
A;Reference number: I59442; MUID:95320273; PMID:7597135
A;Accession: I80329
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-220 <RES>
A;Cross-references: GB:S77585; NID:G957332; PIDN:AAB34728.1; PID:G957333
C;Genetics:
A;Gene: Ptk-3L
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol

Query Match 13.9%; Score 74; DB 2; Length 220;
Best Local Similarity 29.6%; Pred. No. 2.9;
Matches 29; Conservative 9; Mismatches 28; Indels 32; Gaps 6;

QY 1 PGRPTRK-APSHAPLL--GLA-----LLEMHQRHPRRCPPCLCVAGILACGFL 47
DB 80 PYQEPFRGPTHTSAPCPVNGSALLSNPAYVLLATYARPPRGPGPP-----127

QY 48 LGCWG-PSHFQOQCLQALEPOAVSSYLSFGAPLKGRPP 84
DB 128 TPAWAKPTNTQACSGDYMEPE-----KPGAPLLPPPP 159

RESULT 4
A53137
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C;Accession: A53137
R;Sanchez, M.P.; Tapley, P.; Saini, S.S.; He, B.; Pulido, D.; Barbacid, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 1819-1823, 1994
A;Title: Multiple tyrosine protein kinases in rat hippocampal neurons: isolation of Ptk-
A;Reference number: A53137; MUID:94173920; PMID:8127887
A;Accession: A53137
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-910 <RES>
A;Cross-references: GB:L26525; NID:G432480; PIDN:AAA21089.1; PID:G432481
C;Genetics:
A;Gene: Ptk-3
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-ter

C;Keywords: ATP
F;31-186/Domain: discoidin I amino-terminal homology <DN1>
F;605-909/Domain: protein kinase homology <KIN>
F;613-621/Region: protein kinase ATP-binding motif

Query Match 13.9%; Score 74; DB 2; Length 910;
Best Local Similarity 29.6%; Pred. No. 11;
Matches 29; Conservative 9; Mismatches 28; Indels 32; Gaps 6;

QY 1 PGRPTRK-APSHAPLL--GLA-----LLEMHQRHPRRCPPCLCVAGILACGFL 47
DB 480 PYQEPFRGPTHTSAPCPVNGSALLSNPAYVLLATYARPPRGPGPP-----527

QY 48 LGCWG-PSHFQOQCLQALEPOAVSSYLSFGAPLKGRPP 84
DB 528 TPAWAKPTNTQACSGDYMEPE-----KPGAPLLPPPP 559

RESULT 5
AG2114
C;Species: Nostoc sp. PCC 7120
C;Status: preliminary; translated from GB/EMBL/DBJ
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2114
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguic
Nakaaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium A
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AG2114
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-647 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA074169.1; PID:gl7131562; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all2470
C;Superfamily: conserved hypothetical protein sll1033; conserved hypothetical protein ;

Query Match 13.8%; Score 73.5; DB 2; Length 647;
Best Local Similarity 34.5%; Pred. No. 8.9;
Matches 29; Conservative 7; Mismatches 23; Indels 25; Gaps 7;

QY 7 PKAPSHAPLLGLALLRMQHRRPRRCPPCLCVAGILA-----CGFLGCGW---GPHFQQ 58
DB 581 PDSPS-PAPI-----KQPKRRQPLLVLGGLALLVGGTIGLGFAMWQLSPQSFOQ 629

QY 59 SCIQALEPOAVSSYLSFGAPLKGR 82
DB 630 MC-QKL-PPSIQYQYCSPP-----RGR 647

RESULT 6
I46053
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 13-Aug-1999
C;Accession: I46053
R;Gupta, V.K.; Berthoud, V.M.; Atal, N.; Jarillo, J.A.; Barrio, L.C.; Beyer, E.C.
Invest. Ophthalmol. Vis. Sci. 35, 3747-3758, 1994
A;Title: Bovine connexin44, a lens gap junction protein: molecular cloning, immunologi
A;Reference number: I46053; MUID:94375220; PMID:8088962
A;Accession: I46053
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-402 <GUP>
A;Cross-references: EMBL:U08213; NID:G469557; PIDN:AAA50954.1; PID:G469558
C;Superfamily: gap junction protein

Query Match 13.5%; Score 72; DB 2; Length 402;
Best Local Similarity 30.0%; Pred. No. 8.1;
Matches 21; Conservative 6; Mismatches 27; Indels 16; Gaps 1;

QY 28 HPRARCPPLCVAGILACGFLGCGWGFHQQSCLOALEPQAVSYLSPGAPLKGRRPSPG 87
DB 299 HFERRRQP-----GPARRQNWANREAEPTSSRKASPPAPTRIQRAPG 342

QY 88 FORQQRQR 97
DB 343 VAPSPFRRR 352

RESULT 7
B86292
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
C:Accession: B86292
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizlar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: B86292
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1366 <STO>
A:Cross-references: GB:AE005172; NID:G9827657; PIDN:AAF82148.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1

Query Match 13.5%; Score 72; DB 2; Length 1366;
Best Local Similarity 26.1%; Pred. No. 25;
Matches 37; Conservative 16; Mismatches 17; Indels 72; Gaps 9;

QY 1 PCRPTPKAPS--HSAPFLGL-----ALRMHQ-----RHRARCPPL 36
DB 307 PGVPTSATPSTVNSAPLQGLHTNQSSPQLSSQTTQSMRLRHQSSMLRHQPSQ---- 362

QY 37 CVAGILACGFLGCGWGFHQQSCLOALEPQAVSYLSPGAPLKGRRP-----SPG 87
DB 363 QASGI-----HQQS---SLPQOIS-----PLQOQPTQLMRQQAANSSG 399

QY 88 FQR-----QFORQR 97
DB 400 IQQKQMGQGHVGDMMQOQHOOR 421

RESULT 8
A57542
p96 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 23-Feb-1996
C:Accession: A57542
R:Yu, X.X.; Yang, W.; Jackowski, S.; Rock, C.O.
J. Biol. Chem. 270, 14184-14191, 1995
A:Title: Cloning of a novel phosphoprotein regulated by colony-stimulating factor 1 shad
A:Reference number: A57542; MUID:95294028; PMID:7775479
A:Accession: A57542
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-837 <XUA>
A:Cross-references: GB:U1869
A:Note: authors translated the codon GGT for residue 789 as Thr
C:Keywords: alternative splicing; phosphoprotein

Query Match 13.4%; Score 71.5; DB 2; Length 837;
Best Local Similarity 28.4%; Pred. No. 18;
Matches 29; Conservative 11; Mismatches 39; Indels 23; Gaps 5;

QY 1 PCRPTPKAPSHSAPLLGLALLRMHQRHPRARCPP-----LCVAGILACGFLGCGW 52
DB 439 PGRGRTAKSSANDLPASDI-----FASPEPARCPQDNLQSRGRTSWISSKAMILLPQWG 493

QY 53 PSHFQSCLOALEPQ-----VSSYLSPGAPLKGRRPS 85
DB 494 P-FVGLGTVPVTPPOAGPWTVPVYSPSTTVVPGAIISGQPPS 534

RESULT 9
A48280
receptor tyrosine kinase - human
C:Species: Homo sapiens (man)
C:Date: 31-May-1996 #sequence_revision 31-May-1996 #text_change 24-Sep-1999
C:Accession: A48280
R:Johnson, J.D.; Edman, J.C.; Rutter, W.J.
Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681, 1993
A:Title: A receptor tyrosine kinase found in breast carcinoma cells has an extracellular
A:Reference number: A48280; MUID:93296201; PMID:8390675
A:Accession: A48280
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-913 <RES>
A:Cross-references: GB:U11315; NID:G403386; PIDN:AAA02866.1; PID:G403387
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-t
C:Keywords: ATP
F:50-185/Domain: discoidin I amino-terminal homology <DN1>
F:508-912/Domain: protein kinase homology <KIN>
F:616-624/Region: protein kinase ATP-binding motif

Query Match 13.3%; Score 71; DB 2; Length 913;
Best Local Similarity 29.8%; Pred. No. 22;
Matches 29; Conservative 8; Mismatches 29; Indels 32; Gaps 6;

QY 1 PCRPTPKA-PSHSAPLL--GLA-----LLRMHQRHPRARCPPLCVAGILACGFL 47
DB 483 PVCEPRGPNPHSAPCPVNGSALLLSNPAYRLLATVAPRPGPGPP----- 530

QY 48 LGCGW-PSHFQSCLOALEPQAVSYLSPGAPLKGRRP 84
DB 531 TPAAKPTNTQAVSGDYMEP-----KPGAPLLPPPP 562

RESULT 10
B75354
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: B75354
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75354
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-135 <WHI>
A:Cross-references: GB:AE002019; GB:AE000513; NID:G6459547; PIDN:AAF11334.1; PID:G6459
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1772
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1772

Query Match 13.2%; Score 70.5; DB 2; Length 135;
Best Local Similarity 25.5%; Pred. No. 4.2;
Matches 24; Conservative 11; Mismatches 30; Indels 29; Gaps 3;

QY 2 GRPTPKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWGFHQQSC 61

RESULT 13

QJ01486

activin receptor II precursor - human

N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.-)

C:Species: Homo sapiens (man)

C>Date: 17-Jul-1992 #sequence_revision 19-Oct-1995 #text_change 21-Jul-2000

C:Accession: QJ0486; S18908; S22345

R:Donaldson, C.J.; Mathews, L.S.; Vale, W.W.

A:Title: Molecular cloning and binding properties of the human type II activin receptor

A:Reference number: QJ0486, MUID:92231944; PMID:1314589

A:Accession: QJ0486

A:Molecule type: mRNA

A:Residues: 1-513 <DB>

A:Cross-references: GBM>M93415; NID:G178049; PIDN:AAA35504.1; PID:G178050

A:Experimental source: testis

R:Geiser, A.G.

submitted to the EMBL Data Library, December 1991

A:Reference number: S18908

A:Accession: S18908

A:Molecule type: mRNA

A:Residues: 1-513 <GEI>

A:Cross-references: EMBL:X62381; NID:G28347; PIDN:CAA44245.1; PID:G28348

A:Experimental source: mammary gland epithelial cell line B5-589

R:Matzuk, M.M.; Bradley, A.

C:Title: Cloning of the human activin receptor cDNA reveals high evolutionary conservation

A:Reference number: S22345; MUID:92182002; PMID:1311955

A:Accession: S22345

A:Molecule type: mRNA

A:Residues: 1-513 <EMBL>

A:Cross-references: EMBL:X63128; NID:G3928172; PIDN:CAA44839.1; PID:G28350

C:Comment: This protein binds activin A.

C:Genetics:

A:Gene: GDB:ACVR2

A:Cross-references: GDB:132411

A:Map position: 11q13-11q13

C:Superfamily: activin receptor II; protein kinase homology

C:Keywords: ATP; glycoprotein; phosphotransferase; receptor; serine/threonine-specific

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-513/Product: activin receptor II #status predicted <MAT>

F:20-138/Domain: extracellular #status predicted <EXT>

F:139-160/Domain: transmembrane #status predicted <TM1>

F:161-513/Domain: intracellular #status predicted <INT>

F:190-486/Domain: protein kinase homology <KIN>

F:199-206/Region: protein kinase ATP-binding motif

F:43,66/Binding site: carboxylate (Asn) (covalent) #status predicted

F:219/Active site: Lys #status predicted

Query Match 13.1%; Score 70; DB 1; Length 513;

Best Local Similarity 27.7%; Pred. No. 16;

Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4;

QY 4 PTPKAPSHSAPLGLALLRMHQRHRCBPPLCVAGILACGFLGCGWSPHFQSCLOA 63

DB 129 PVTPKPPYNYLLISLV-----PLMLIAGIVICAF----WYRRHHKMAYPEV 171

QY 64 LEQAVSSYLSPGAPLGRPPSP 86

DB 172 LVPTQ-----DPGPP-----PESP 185

RESULT 14

A39896

activin receptor precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 18-Jun-1999

C:Accession: A39896

R:Mathews, L.S.; Vale, W.W.

Cell 65, 973-982, 1991

A:Title: Expression cloning of an activin receptor, a predicted transmembrane serine k

A:Reference number: A39896; MUID:91256317; PMID:1646080
A:Accession: A39896
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-513 <MAT>
A:Cross-references: GB:M65287; NID:g191663; PIDN:AAA37171.1; PID:g191664
C:Superfamily: activin receptor II; protein kinase homology
C:Keywords: ATP; receptor; serine/threonine-specific protein kinase; transmembrane prote
F:190-486/Domain: protein kinase homology <KIN>

Query Match 13.1%; Score 70; DB 2; Length 513;
Best Local Similarity 27.7%; Pred. No. 16;
Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4;
QY 4 PTPKAPSHSAPILGLALLRMHQHPRARCPLCVAGILACGFLGCGWGPSHFQQSCLOA 63
Db 129 PVTPKPPYNYLLYSLV-----PLMLIAGIVICAF-----WYRHHKWAYPPV 171

QY 64 LEQAVSSYLSGAPLKGPPSP 86
Db 172 LVPTQ-----DPGPP-----PPSP 185

RESULT 15
A49193
type II activin receptor ActRII - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 23-May-1997
C:Accession: A49193
R:Feng, Z.M.; Madigan, M.B.; Chen, C.L.
Endocrinology 132, 2593-2600, 1993
A:Title: Expression of type II activin receptor genes in the male and female reproductiv
A:Reference number: A49193; MUID:93279247; PMID:7916681
A:Accession: A49193
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-513 <FEN>
A>Note: sequence extracted from NCBI backbone (NCBIN:133008, NCBIP:133009)
C:Superfamily: activin receptor II; protein kinase homology
C:Keywords: ATP; receptor
F:190-486/Domain: protein kinase homology <KIN>

Query Match 13.1%; Score 70; DB 2; Length 513;
Best Local Similarity 27.7%; Pred. No. 16;
Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4;
QY 4 PTPKAPSHSAPILGLALLRMHQHPRARCPLCVAGILACGFLGCGWGPSHFQQSCLOA 63
Db 129 PVTPKPPYNYLLYSLV-----PLMLIAGIVICAF-----WYRHHKWAYPPV 171

QY 64 LEQAVSSYLSGAPLKGPPSP 86
Db 172 LVPTQ-----DPGPP-----PPSP 185

Search completed: March 13, 2004, 07:44:38
Job time : 5.3849 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 2.02083 Seconds
(without alignments)
2499.367 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534

Sequence: 1 FGRPRPKAPSHSFLGLA.....PLKGRPPSPGFORQRR 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	79.5	14.9	353	1 ZIC4 MOUSE	Q61467 mus musculus
2	75	14.0	958	1 TRH DROME	Q24119 drosophila
3	74.5	14.0	2337	1 TOR2 SCHPO	Q9V7K2 schizosacch
4	74	13.9	910	1 DDR1 RAT	Q63474 rattus norv
5	72	13.5	325	1 NPDI HUMAN	Q9NQX5 homo sapien
6	72	13.5	911	1 DDR1 MOUSE	Q03146 mus musculus
7	71	13.3	513	1 AVR2 SHEEP	Q28560 ovis aries
8	71	13.3	909	1 DDR1 PANTR	Q7YR43 pan troglod
9	71	13.3	913	1 DDR1 HUMAN	Q08345 h epithelia
10	70.5	13.2	274	1 THA8 HUMAN	Q8NA92 homo sapien
11	70.5	13.2	606	1 NQOC THETH	Q56227 thermus the
12	70	13.1	513	1 AVR2 BOVIN	Q28043 bos taurus
13	70	13.1	513	1 AVR2 HUMAN	P27037 homo sapien
14	70	13.1	513	1 AVR2 MOUSE	P27038 mus musculus
15	70	13.1	562	1 AT15 MOUSE	P59384 mus musculus
16	70	13.1	1083	1 T2D3 HUMAN	O00268 homo sapien
17	69	12.9	258	1 IBP4 HUMAN	P22692 homo sapien
18	69	12.9	513	1 AVR2 RAT	P38444 rattus norv
19	68	12.7	550	1 MYCO STRCI	P20910 streptomyc
20	68	12.7	1042	1 SPAL HUMAN	Q95F54 homo sapien
21	68	12.7	3703	1 ABF1 HUMAN	Q15911 homo sapien
22	67.5	12.6	239	1 NPDI HUMAN	Q15522 homo sapien
23	67.5	12.6	332	1 KPD1 MOUSE	Q64322 mus musculus
24	67.5	12.6	618	1 NM24 MOUSE	Q9T0S2 mus musculus
25	67.5	12.6	684	1 EP84 HCMVA	P17151 human cytom
26	67.5	12.6	955	1 B3A4 RABIT	Q9SKY1 oryctolagus
27	67	12.5	2167	1 SHK1 RAT	Q9WV48 rattus norv
28	67	12.5	5262	1 MLL2 HUMAN	Q14686 homo sapien
29	66.5	12.5	442	1 ETBR MOUSE	Q49302 mus musculus
30	66.5	12.5	589	1 AXUL HUMAN	Q9SE65 homo sapien
31	66.5	12.5	772	1 Z469 HUMAN	Q96J99 homo sapien
32	66	12.4	116	1 STP2 RAT	P1101 rattus norv
33	66	12.4	245	1 CM32 HUMAN	Q8ND00 homo sapien

ALIGNMENTS

RESULT 1				
ZIC4 MOUSE				
ID	ZIC4 MOUSE	STANDARD;	PRT;	353 AA.
AC	Q61467;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Zinc finger protein ZIC 4 (Zinc finger protein of the cerebellum 4).			
GN	ZIC4.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cerebellum;			
RX	MEDLINE=96269420; PubMed=8682319;			
RA	Aruga J., Yozu A., Hayashizaki Y., Okazaki Y., Chapman V.M.,			
RA	Mikoshiba K.;			
RT	"Identification and characterization of Zic4, a new member of the			
RT	mouse Zic gene family";			
RL	Gene 172:291-294(1996).			
CC	-!- SUBCELLULAR LOCATION: Nuclear.			
CC	-!- TISSUE SPECIFICITY: CNS. A high level expression is seen in the			
CC	cerebellum.			
CC	-!- SIMILARITY: BELONGS TO THE GLI FAMILY OF C2H2-TYPE ZINC-FINGER			
CC	PROTEINS.			
CC	-!- SIMILARITY: Contains 4 C2H2-type zinc fingers.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; D78174; BAA11240.1; --			
DR	PIR; J4875; J4875.			
DR	HSSP; P08047; ISP2.			
DR	MGD; MGI:107201; Zic4.			
DR	InterPro; IPR007087; Znf C2H2.			
DR	Pfam; PF00096; zf-C2H2; 4.			
DR	ProDom; PD000003; Znf C2H2; 1.			
DR	SMART; SM00355; Znf C2H2; 4.			
DR	PROSITE; PS00028; ZINC FINGER C2H2_1; 3.			
DR	PROSITE; PS00157; ZINC FINGER C2H2_2; 4.			
KW	Zinc-finger; Metal-binding; DNA-binding; Repeat; Nuclear protein.			
FT	ZN_FING 183 205 C2H2-TYPE 1 (DEGENERATE).			
FT	ZN_FING 211 235 C2H2-TYPE 2.			
FT	ZN_FING 241 265 C2H2-TYPE 3.			
FT	ZN_FING 271 295 C2H2-TYPE 4.			
SQ	SEQUENCE 353 AA; 38692 MW; CDB1F390884793C2 CRC64;			
Query Match 14.9%; Score 79.5; DB 1; Length 353;				
Best Local Similarity 27.9%; Pred. No. 1.1;				

P20186 streptomyc
Q8IN94 drosophila
Q9Y26 homo sapien
P21860 homo sapien
P40226 mus musculus
O75038 homo sapien
Q02078 homo sapien
P31094 synchococ
O75385 homo sapien
Q9C0C2 homo sapien
Q9C203 mus musculus
P21917 homo sapien

Matches 31; Conservative 12; Mismatches 49; Indels 19; Gaps 4;

QY 1 PGRPRKAPKSPAPLLGLALL---RHHORH-----PRARPPLCVAGILACGF----- 46

DB 51 POLHQQPPQASHSRPLNGLRLIGIPGDMYARSPFPAGPWARSDTLATATLH-GYGGMN 109

QY 47 ---LLGCWGPSPHFQSCQLQALPQAVSSYISFGAPLKGSPSPGFORQQRQ 93

DB 110 LTMNLTAHPGCAFFRYMQPIKQELICKLWLGDDSPMSRPFCSKTFSTMHE 160

RESULT 2

TRH_DROME STANDARD; PRT; 958 AA.

AC Q24119; Q24165; Q8SX13; Q9WQ07;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 15-MAR-2004 (Rel. 43, Last annotation update)

DE Trachealeless protein.

GN TRH OR CG6883

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID:7227;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).

RC TISSUE=Embryo;

RX MEDLINE=96136712; PubMed=8557198;

RA Wilk R., Weizman I., Shilo B.-Z.;

RT "Trachealeless encodes a bHLH-PAS protein that is an inducer of

RT tracheal cell fates in Drosophila.";

RL Genes Dev. 10:93-102(1996).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Embryo;

RX MEDLINE=96136713; PubMed=8557189;

RA Isaac D.D., Andrew D.J.;

RT "Tubulogenesis in Drosophila: a requirement for the trachealeless gene

RT product.";

RL Genes Dev. 10:103-117(1996).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.P.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brotter P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablo J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,

RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,

RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

RN [4]

RP SEQUENCE OF 77-958 FROM N.A. (ISOFORM 1).

RC STRAIN=Berkley; TISSUE=Head;

RX MEDLINE=22426066; PubMed=12537569;

RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,

RA George R.A., Guarin H., Kronmiller B., Pacle J.M., Park S., Wan K.H.,

RA Rubin G.M., Celniker S.E.; cDNA resource.";

RT "A Drosophila full-length cDNA resource.";

RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).

CC -!- FUNCTION: Transcription factor, master regulator of tracheal cell

CC fates in the embryo, necessary for the development of the salivary

CC gland duct and the posterior spiracles. It may induce a general

CC rate of branched tubular structures of epithelial origin. TGO/TRH

CC heterodimers are involved in the control of breathless expression.

CC -!- SUBUNIT: Efficient DNA binding requires dimerization with another

CC bHLH protein. Heterodimer with TGO.

CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS:

CC -!- Event=Alternative splicing; Named isoforms=3;

CC Comment=Additional isoforms seem to exist;

CC Name=1; Synonyms=A;

CC IsoId=Q24119-1; Sequence=Displayed;

CC Name=2;

CC IsoId=Q24119-2; Sequence=VSP_002161;

CC Name=3;

CC IsoId=Q24119-3; Sequence=VSP_002162;

CC -!- TISSUE SPECIFICITY: Trachea, salivary gland ducts, posterior

CC spiracles (Filzkörper primordia) and a subset of cells in the CNS.

CC -!- DEVELOPMENTAL STAGE: During embryogenesis, first detected in the

CC tracheal placodes at stage 8, and expression continues throughout

CC embryonic and larval development. In the developing salivary

CC gland, expression is observed in the entire gland at stage 9 and

CC by stage 12, expression is confined to the salivary ducts.

CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.

CC -!- SIMILARITY: Contains 2 PAS (PER-ARNT-SYM) dimerization domains.

CC -!- SIMILARITY: Contains 1 PAS-associated C-terminal (PAC) domain.

CC -!- CAUTION: Ref.2 sequence differs from that shown due to frameshifts

CC in positions 22 and 34.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; U33427; AAA96257.1; ALT INIT.

CC EMBL; U42699; AAA96754.1; ALT_FRAME.

CC EMBL; AE003468; AAF47386.1;

CC EMBL; AY094911; AAM11264.1; ALT_INIT.

CC FlyBase; FBgn0003749; trh.

CC GO; GO:0005634; C:nucleus; ISS.

CC GO; GO:0003702; F:RNA polymerase II transcription factor acti...; ISS.

CC GO; GO:0006355; P:regulation of transcription, DNA-dependent; ISS.

CC GO; GO:0007431; P:salivary gland development; IMP.

CC GO; GO:0007424; P:tracheal system development (sensu Insecta); IMP.

CC InterPro; IPR001092; HLH_basic.

CC InterPro; IPR001610; PAC.

CC InterPro; IPR000014; PAS_domain.

CC Pfam; PF00010; HLH; 1.

CC Pfam; PF00989; PAS; 1.

CC SMART; SM00353; HLH; 1.

DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00888; HLH; 1.
DR PROSITE; PS0112; PAS; 2.
KW Developmental protein; Nuclear protein; Transcription regulation;
Repeat; DNA-binding; Alternative splicing.
FT DNA_BIND 86 99
FT DOMAIN 100 140 BASIC DOMAIN.
FT DOMAIN 174 244 HELIX-LOOP-HELIX MOTIF.
FT DOMAIN 391 461 PAS 1.
FT DOMAIN 465 508 PAS 2.
FT DOMAIN 629 636 PAC.
FT DOMAIN 154 157 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT DOMAIN 244 250 POLY-SER.
FT DOMAIN 251 254 POLY-GLY.
FT DOMAIN 255 264 POLY-SER.
FT DOMAIN 722 728 POLY-GLY.
FT VARSPLIC 251 286
FT VARSPLIC 328 356 Missing (in isoform 2).
FT VARSPLIC 328 356 Missing (in isoform 3).
FT VARSPLIC 328 356 /FTID=VSP_002161.
FT VARSPLIC 328 356 /FTID=VSP_002162.
FT VARSPLIC 328 356 Missing (in isoform 1).
FT VARSPLIC 328 356 P -> A (IN REF. 1).
FT VARSPLIC 328 356 G -> GG (IN REF. 2).
FT VARSPLIC 328 356 A -> T (IN REF. 1).
FT VARSPLIC 328 356 A -> F (IN REF. 1).
FT VARSPLIC 328 356 A -> V (IN REF. 1).
SQ SEQUENCE 958 AA; 102241 MW; 8F9CF758F1370541 CRC64;

Query Match 14.0%; Score 75; DB 1; Length 958;
Best Local Similarity 26.8%; Pred. No. 7.6;
Matches 30; Conservative 15; Mismatches 35; Indels 32; Gaps 5;

Qy 3 RTRPKAPSHS-----APLLGLALLRMVHQRHPRACPLVAGILACGFLLCGW 51
Db 630 RRRKTKASQAEDEQGEQVISEQPLKPLTMEQRDQPRSLPSIVDE----- 677

Qy 52 GPSHFQSCLOALEPOAVSSYL-SPGAPLGRPPSPGF-----QRQRQ 95
Db 678 QFSSAADSAVKLE-QAMSKHLPSPAAVSVAPNTDFADSLLKQQQQQQ 728

RESULT 3
TOR2_SCHPO STANDARD; PRT; 2337 AA.
ID TOR2_SCHPO
AC Q9Y7X2; O94507;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 3-kinase tor2 (EC 2.7.1.137) (PI3-kinase)
DE (PtdIns-3-kinase) (PI3K).
GN TOR2 OR SPBC216.07C OR SPBC646.01C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RC MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hanlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson S., Quail M.A., Rabinowitz E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,

Query Match 14.0%; Score 74.5; DB 1; Length 2337;
Best Local Similarity 29.9%; Pred. No. 20;
Matches 20; Conservative 10; Mismatches 28; Indels 9; Gaps 3;

QY 17 LGALLRMRHPRARCPPLCVIA-----GILAGFLLGWGP--SHFQQSCLOALEPQA 68
D5 1137 LGVALLRMRHPRARCPPLCVIA-----GILAGFLLGWGP--SHFQQSCLOALEPQA 68

QY 69 VSSYLSF 75
D5 1196 TSFHSF 1202

RESULT 4

DDRI RAT STANDARD; PRT; 910 AA.

AC Q63474;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Epithelial discoidin domain receptor 1 precursor (EC 2.7.1.112)
DE (Tyrosine kinase DDR) (Discoidin receptor tyrosine kinase) (Tyrosine-
protein kinase CAK) (Cell adhesion kinase) (Protein-tyrosine kinase
PTK-3)
GN DDRI OR EDDR1 OR PTK3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; TISSUE=Brain;
RX MEDLINE=94173920; PubMed=8127887;
RA Sanchez M.P., Tapley P., Saini S.S., He B., Pulido D., Barbacid M.;
RT "Multiple tyrosine protein kinases in rat hippocampal neurons: Isolation of Ptk-3, a receptor expressed in proliferative zones of the developing brain";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1819-1823(1994).
CC -!- FUNCTION: May be involved in cell-cell interactions and recognition (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Various embryonic and adult tissues; also proliferative zones of the developing brain; hippocampal neurons.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin receptor subfamily.
CC -!- SIMILARITY: Contains 1 FS/8 type C domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L26525; AAA21089.1; -
CC PIR; A53137; A53137.
CC HSPSP; P00523; 2PTK.
CC InterPro; IPR000421; FA58 C.
CC InterPro; IPR008979; Gal_Eind_like.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR002011; RecepttyrkinII.
CC InterPro; IPR001245; Tyr_pkinase.
CC InterPro; IPR008266; Tyr_pkinase_AS.
CC Pfam; PF00754; FS_F8_type_C; 1.
CC Pfam; PF00069; Pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SMC0231; FA58C; 1.
CC SMART; SMC0219; TyrKs; 1.
CC PROSITE; PS01285; FA58C_1; 1.

DR PROSITE; PS01286; FA58C_2; 1.
DR PROSITE; PS00022; FA58C_3; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
KW Transferase; Tyrosine-protein kinase; Glycoprotein; Signal;
KW Phosphorylation; Transmembrane; Receptor; ATP-binding.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 910 EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1.
FT DOMAIN 20 413 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 414 440 POTENTIAL.
FT DOMAIN 441 910 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 32 186 FS/8 TYPE C.
FT DOMAIN 378 412 GLY/PRO-RICH.
FT DOMAIN 473 598 GLY/PRO-RICH.
FT DOMAIN 607 902 PROTEIN KINASE.
FT NP BIND 613 621 ATP (BY SIMILARITY).
FT BINDING 652 652 ATP (BY SIMILARITY).
FT ACT_SITE 763 763 BY SIMILARITY.
FT DISULFID 32 186 BY SIMILARITY.
FT MOD_RES 510 510 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 789 789 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 793 793 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 794 794 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 212 212 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 261 261 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 371 371 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 910 AA; 101166 MW; 7E7FFAIDCB029806 CRC64;
Query Match 13.9%; Score 74; DB 1; Length 910;
Best Local Similarity 29.6%; Pred. No. 9;
Matches 29; Conservative 9; Mismatches 28; Indels 32; Gaps 6;

QY 1 FGRPTPK-APSHAPLL--GLA-----LLRMHQRHPRARCPPLCVAGILACGFL 47
D5 480 PYQEPRTGPTTHSAPCVNGSALLSNPAYRLLATVAPRPGPGPP----- 527

QY 48 LGCWG-PSHFQSCLOALEPQAVSSYLSFGAPLKRPP 84
D5 528 TPAWAKPTNTQACSGDYWEPE-----KPGAPLLPPPP 559

RESULT 5

NPDI HUMAN STANDARD; PRT; 325 AA.

ID NPDI HUMAN Q9NQX5; Q9BTD6; Q9EXT3; Q9NQS2; Q9Y434;
AC Q9NQX5; Q9BTD6; Q9EXT3; Q9NQS2; Q9Y434;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Neural proliferation differentiation and control protein-1 precursor (NPDC-1 protein).
GN NPDC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP TISSUE=Fetal liver;
RX MEDLINE=21142391; PubMed=11245976;
RA Qu X., Zhang C., Zhai Y., Xing G., Wei H., Yu Y., Wu S., He F.;
RT "Characterization and tissue expression of a novel human gene npdc1.";
RL Gene 264:37-44(2001).
RN [2]
SEQUENCE FROM N.A.
RP TISSUE=I-tis;
RC TISSUE=I-tis;
RA Wistow G.J.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;


```

DR InterPro: IPR000719; Prot kinase.
DR InterPro: IPR002011; RecepttyrkinasII.
DR InterPro: IPR001245; Tyr_kinase_II.
DR InterPro: IPR008266; Tyr_pkinase_AS.
DR Pfam: PF00754; F5_F8_type_C; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Prot_kinase; 1.
DR SMART: SM00231; FA58C_1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS01285; FA58C_1; 1.
DR PROSITE: PS01286; FA58C_2; 1.
DR PROSITE: PS50022; FA58C_3; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00239; RECEPTOR_TYR_KIN_II; 1.
DR Transferrase; Tyrosine-protein Kinase; Glycoprotein; Signal;
KW Phosphorylation; Transmembrane; Receptor; ATP-binding;
KW Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 911
FT DOMAIN 20 414
FT TRANSMEM 415 441
FT DOMAIN 442 911
FT DOMAIN 32 186
FT DOMAIN 379 413
FT DOMAIN 474 599
FT DOMAIN 608 903
FT NP_BIND 614 622
FT BINDING 653 653
FT ACT_SITE 764 764
FT DISULFID 32 186
FT MOD_RES 511 511
FT MOD_RES 790 790
FT MOD_RES 794 794
FT MOD_RES 795 795
FT CARBOHYD 213 213
FT CARBOHYD 262 262
FT CARBOHYD 372 372
FT CARBOHYD 392 392
FT VARSPIC 503 539
SQ SEQUENCE 911 AA; 101160 MW; DBB7FE03DD79510 CRC64;

Query Match 13.5%; Score 72; DB 1; Length 911;
Best Local Similarity 29.6%; Pred. No. 14;
Matches 29; Conservative 8; Mismatches 29; Indels 32; Gaps 6;

QY 1 PGRTPRK-APSHSAPLL--GLA-----LLRMHQRHPRARCPPLCVAGILACQFL 47
DB 481 PYQEPGRGTPPHSAPCVENGSAALLSNPAYRLLLATVAPRPGPGPP-----528

QY 48 LGCWG-PSHFQQSCLQALEPOAVSSYLSPGAPLKRPP 84
DB 529 TPWAKPTNTQACSGDYMEP-----KPGAPLLPPPP 560

RESULT 7
AVR2_SHEEP STANDARD; PRT; 513 AA.
AC Q28560;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Activin receptor type II precursor (EC 2.7.1.37) (ACTR-II).
GN ACVR2 OR ACTRII.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-Romney; TISSUE=Ovarian follicle;
RA Tisdall D.J.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Receptor for activin A, activin B and inhibin A.
CC -!- Involved in transmembrane signaling.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with AIP1. Part of a complex consisting of
CC AIP1, ACVR2, ACVR1B and MADH3 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. TGFB
CC receptor subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L19442; AAA91903.1; -.
CC HSP: P27038; 187E.
DR InterPro: IPR000472; Activin receptor.
DR InterPro: IPR000333; Actn_receptorII.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_pkin_AS.
DR Pfam: PF01064; Activin_recp; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00653; ACTIVIN2R.
DR ProDom: PD000001; Prot_kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
KW Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 19
FT CHAIN 20 513
FT DOMAIN 20 135
FT TRANSMEM 136 161
FT DOMAIN 162 513
FT DOMAIN 192 485
FT NP_BIND 198 206
FT BINDING 219 219
FT ACT_SITE 322 322
FT DISULFID 30 60
FT DISULFID 50 78
FT DISULFID 85 104
FT DISULFID 91 103
FT DISULFID 105 110
FT CARBOHYD 43 43
FT CARBOHYD 66 66
FT CARBOHYD 513 AA; 57768 MW; 7231BF9E85CA57E3 CRC64;
SQ SEQUENCE 513 AA; 57768 MW; 7231BF9E85CA57E3 CRC64;

Query Match 13.3%; Score 71; DB 1; Length 513;
Best Local Similarity 28.9%; Pred. No. 10;
Matches 24; Conservative 6; Mismatches 27; Indels 26; Gaps 4;

QY 4 PTPKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACQFLGCGWPSHFQQSCLQA 63
DB 129 PVTPKPPYNYLLYSLV-----PLMLVAGIVICAF-----WYRHKKWAPPV 171

QY 64 LEQAVSSYLSPGAPLKRPPSP 86
DB 172 LVPTQ-----DPGPP-----PPSP 185

RESULT 8
DDR1_PANTR STANDARD; PRT; 909 AA.
ID DDR1_PANTR
AC QYR43;
DT 15-MAR-2004 (Rel. 43, Created)
DT 15-MAR-2004 (Rel. 43, Last sequence update)

```

15-MAR-2004 (Rel. 43, Last annotation update)
Epithelial discoidin domain receptor 1 precursor (EC 2.7.1.112)
(Tyrosine kinase DDR) (Discoidin receptor tyrosine kinase).
DDRI OR EDDR1
Pan troglodytes (Chimpanzee).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
NCBI TaxID=9598;
[1]
SEQUENCE FROM N.A.
MEDLINE=22709134; PubMed=12799463;
Anzai T., Shiina T., Kimura N., Yanagiya K., Kohara S., Shigenari A.,
Yamagata T., Kulski J.K., Naruse T.K., Fujimori Y., Fukuzumi Y.,
Yamazaki M., Tashiro H., Iwanoto C., Umebara Y., Imanishi T.,
Meyer A., Ikeo K., Gotohori T., Bahram S., Inoko H.;
"Comparative sequencing of human and chimpanzee MHC class I regions
unveils insertions/deletions as the major path to genomic
divergence.";
Proc. Natl. Acad. Sci. U.S.A. 100:7708-7713 (2003).
-!- FUNCTION: May be involved in cell-cell interactions and
recognition (by similarity).
-!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
-!- SUBCELLULAR LOCATION: Type I membrane protein.
-!- DOMAIN: The Gly/Pro-rich domains may be required for an unusual
geometry of interaction with ligand or substrates.
-!- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
receptor subfamily.
-!- SIMILARITY: Contains 1 F5/8 type C domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@sib-sib.ch).
EMBL; ABL00084; BAC78172.1; -.
InterPro; IPR000421; FA58_C.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR002011; RecepttyrklnsII.
InterPro; IPR002290; Ser_thr_kinase.
InterPro; IPR001245; Tyr_kinase.
InterPro; IPR008266; Tyr_kinase_AS.
Pfam; PF00754; F5_F8_type_C; 1.
Pfam; PF00069; pkinase; 1.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00231; FA58C; 1.
SMART; SM00220; S_TKC; 1.
SMART; SM00219; TyrKC; 1.
PROSITE; PS01285; FA58C_1; 1.
PROSITE; PS01286; FA58C_2; 1.
PROSITE; PS00022; FA58C_3; 1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
Transferase: Tyrosine-protein kinase; Glycoprotein; Signal;
Phosphorylation; Transmembrane; receptor; ATP-binding.
FT SIGNAL 1 18
FT CHAIN 19 909
FT DOMAIN 19 412
FT TRANSMEM 413 439
FT DOMAIN 440 909
FT DOMAIN 31 181
FT DOMAIN 373 411
FT DOMAIN 472 597
FT DOMAIN 606 901
FT NP_BIND 612 620
FT BINDING 651 651
FT ACT_SITE 762 762
FT DISULFID 31 181

FT MOD_RES 509 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 788 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 792 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES 793 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CARBOHYD 207 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 256 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 366 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 390 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 909 AA; 100641 MW; 5E30275EE6D0D931 CRC64;
Query Match 13.3%; Score 71; DB 1; Length 909;
Best Local Similarity 29.6%; Pred. No. 18;
Matches 29; Conservative 8; Mismatches 29; Indels 32; Gaps 6;
QY 1 PGRTPRPA-PHSAPLL--GLA-----LLRMHQHPRARCPPLCVAGILACGFL 47
Db 479 PYQEPGRGPNPHSAFCVNGSALLLSNPAYRLLATYARPPRGPGPP----- 526
QY 48 LGCWG-PSHFQSCLOALPQAVSSVLSFGAPLKGRRP 84
Db 527 TPAWAKPTNTQAYSGDYMEPE-----KPGAPLLPFP 558
RESULT 9
DDRI_HUMAN
ID DDRI_HUMAN STANDARD; PRT; 913 AA.
AC Q08345; Q14196; Q16562; Q9UD37;
DT 01-FEB-1995 (Rel. 31, Created).
DT 01-NOV-1997 (Rel. 35, Last sequence update).
DT 10-OCT-2003 (Rel. 42, Last annotation update).
DE Epithelial discoidin domain receptor 1 precursor (EC 2.7.1.112)
DE (Tyrosine kinase DDR) (Discoidin receptor tyrosine kinase) (Tyrosine-
protein kinase CAK) (Cell adhesion kinase) (TRK E) (protein-tyrosine
kinase RTK 6) (CD167a antigen) (HGK2).
DE DDRI OR EDDR1 OR CAK OR TRKE OR RTK6.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=94134417; PubMed=8302582;
RA Perez J.L., Shen X., Finkernagel S., Sciorra L., Jenkins N.A.,
Gilbert D.J., Copeland N.G., Wong T.W.;
"Identification and chromosomal mapping of a receptor tyrosine kinase
with a putative phospholipid binding sequence in its ectodomain.";
Oncogene 9:211-219 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93296201; PubMed=8390675;
RA Johnson J.D., Edman J.C., Rutter W.J.;
"A receptor tyrosine kinase found in breast carcinoma cells has an
extracellular discoidin I-like domain.";
Proc. Natl. Acad. Sci. U.S.A. 90:5677-5681 (1993).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97131588; PubMed=8977099;
RA Sakuma S., Tada M., Sava H., Sawamura Y., Shinoh Y., Abe H.;
"Receptor protein tyrosine kinase DDR is up-regulated by p53
protein.";
FEBS Lett. 398:165-169 (1996).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96389017; PubMed=8796349;
RA Playford M.P., Butler R.J., Wang X.C., Katsco R.M., Cooke I.E.,
Gareean T.S.;
"The genomic structure of discoidin receptor tyrosine kinase.";
Genome Res. 6:620-627 (1996).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Lung;

RX MEDLINE=96204002; PubMed=8622863;
 RA Perez J.L., Jing S.Q., Wong T.W.;
 RT "Identification of two isoforms of the Cak receptor kinase that are
 RL coexpressed in breast tumor cell lines.";
 RN Oncogene 12:1469-1477(1996).
 RP [6]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Brain, and Keratinocytes;
 RX MEDLINE=94043265; PubMed=8226977;
 RA di Marco E., Cutuli N., Guerra L., Cancedda R., de Luca M.;
 RT "Molecular cloning of trkE, a novel trk-related putative tyrosine
 RL kinase receptor isolated from normal human keratinocytes and widely
 RN expressed by normal human tissues.";
 RP J. Biol. Chem. 268:24290-24295(1993).
 RP [7]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Ovary;
 RX MEDLINE=95151638; PubMed=7843919;
 RA Laval S., Butler R., Shelling A.N., Hanby A.M., Poulson R.,
 RP Ganesan T.S.;
 RT "Isolation and characterization of an epithelial-specific receptor
 RL tyrosine kinase from an ovarian cancer cell line.";
 RN Cell Growth Differ. 5:1173-1183(1994).
 RP [8]
 RP SEQUENCE FROM N.A.
 RA Shina S., Tamiya G., Oka A., Inoko H.;
 RT "Homo sapiens 2,229,817bp genomic DNA of sp21.3 HLA class I region.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=Muscle;
 RX MEDLINE=22398257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunatratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Heiton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.W., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP [10]
 RP SEQUENCE OF 771-824 FROM N.A.
 RX MEDLINE=95136019; PubMed=7834423;
 RA Weiner T.M., Liu E.T., Craven R.J., Cance W.G.;
 RT "Expression of growth factor receptors, the focal adhesion kinase, and
 RL other tyrosine kinases in human soft tissue tumors.";
 RN Ann. Surg. Oncol. 1:18-27(1994).
 CC -1- FUNCTION: May be involved in cell-cell interactions and
 CC recognition.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1; Synonyms=CAK I;
 CC IsoId=Q08345-1; Sequence=Displayed;
 CC Name=2; Synonyms=CAK II, Short;
 CC IsoId=Q08345-2; Sequence=VSP_002953;
 CC -1- TISSUE SPECIFICITY: Expressed at low levels in most adult tissues
 CC and is highest in the brain and lung. Abundant in breast carcinoma
 CC cell lines.

CC -1- DOMAIN: The Gly/Pro-rich domains may be required for an unusual
 CC geometry of interaction with ligand or substrates.
 CC -1- SIMILARITY: Belongs to the Tyr family of protein kinases. Insulin
 CC receptor subfamily.
 CC -1- SIMILARITY: Contains 1 F5/8 type C domain.
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1986613159_g.htm".
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L11315; AAA02866.1; -;
 DR EMBL; L20817; AAA18019.1; -;
 DR EMBL; U48705; AAC50917.1; -;
 DR EMBL; X98208; CAA66871.1; -;
 DR EMBL; X99023; CAA66871.1; JOINED.
 DR EMBL; X99024; CAA66871.1; JOINED.
 DR EMBL; X99025; CAA66871.1; JOINED.
 DR EMBL; X99026; CAA66871.1; JOINED.
 DR EMBL; X99027; CAA66871.1; JOINED.
 DR EMBL; X99028; CAA66871.1; JOINED.
 DR EMBL; X99029; CAA66871.1; JOINED.
 DR EMBL; X99030; CAA66871.1; JOINED.
 DR EMBL; X99031; CAA66871.1; JOINED.
 DR EMBL; X99032; CAA66871.1; JOINED.
 DR EMBL; X99033; CAA66871.1; JOINED.
 DR EMBL; X99034; CAA66871.1; JOINED.
 DR EMBL; L57508; AAB05208.1; -;
 DR EMBL; X74979; CAA52915.1; -;
 DR EMBL; Z29093; CAA82335.1; -;
 DR EMBL; AP000511; BAB63318.1; -;
 DR EMBL; BC008716; AAH08716.1; -;
 DR EMBL; BC013400; AAH13400.1; -;
 DR F1R; A48280; A48280.
 DR F1R; A49508; A49508.
 DR HSP; P11362; IFGK.
 DR Genew; HGNC:2730; DDR1.
 DR MIM; 600408; -;
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:000714; F:transmembrane receptor protein tyrosine kin. . . ; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR InterPro; IPR000421; FA58_C.
 DR InterPro; IPR008979; Gal_Bind_like.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR02011; RecepttyrkinasII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR InterPro; IPR008266; Tyr_pkinase_AS.
 DR Pfam; PF00754; F5_F8_type_C; 1.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00231; FA58C; 1.
 DR SMART; SM00219; TyrKc; 1.
 DR PROSITE; PS01285; FA58C_1; 1.
 DR PROSITE; PS01285; FA58C_2; 1.
 DR PROSITE; PS00022; FA58C_3; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00107; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
 DR TRANSFERASE; Tyrosine-protein Kinase; Glycoprotein; Signal;
 KW Phosphorylation; Transmembrane; Receptor; ATP-binding; Antigen;
 KW Alternative splicing.
 FT SIGNAL 1 18
 FT CHAIN 19 913
 FT DOMAIN 19 416
 FT TRANSMEM 417 443
 FT DOMAIN 444 913

FT DOMAIN 31 185 F5/8 TYPE C. 82
FT DOMAIN 377 415 GLY/PRO-RICH.
FT DOMAIN 476 601 GLY/PRO-RICH.
FT DOMAIN 610 905 PROTEIN KINASE.
FT NP BIND 616 624 ATP (BY SIMILARITY).
FT BINDING 655 655 ATP (BY SIMILARITY).

Query Match 13.3%; Score 71; DB 1; Length 913;
Best Local Similarity 29.6%; Pred. No. 18;
Matches 29; Conservative 8; Mismatches 29; Indels 32; Gaps 6;

QY 1 PGPRTRPKA-PSHAPLL-GLA-----LLRMEQRHPRARCPPLCVAGILACGFL 47
DB 483 FYQEPRGRGNPPHSPAPCVNGSALLSNPFAVRLLLATVARPRGPGPP----- 530

QY 48 LGCWG-PSHQOQCLQALEPQAVSSYLSGAPLKCRPP 84
DB 531 TPAWAKPTNTQAYSGDYMEPE-----XPGAPLLPFP 562

RESULT 10
THA8_HUMAN STANDARD; PRT; 274 AA.
AC Q8N932; Q96W21;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE THAP domain protein 8.
GN THAP8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,
RA Hotaue T., Hiraoka S., Murakawa K., Takiguchi S., Kusano J., Chiba Y.,
RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saico K., Nishikawa T.,
RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Takatsuka M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Wazuki Y., Sugano S., Naganari K., Masuo Y., Nagai K., Isogai T.;
RA "NEDO human cDNA sequencing project";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Contains 1 THAP domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AK057453; BAB71493.1; -
CC EMBL; AK093048; BAC04034.1; -
CC Genew; HGNC:23191; THAP8.
CC InterPro; IPR008612; DUF_DM3.
CC Pfam; PF05485; THAP; 1.
CC SMART; SM00692; DM3; 1.
CC Zinc-finger; DNA-binding.
CC
CC DOMAIN 1 90 THAP.
FT ZN FING 5 61 THAP-TYPE.
FT ZN FING 185 185 R -> Q (IN REF. 1; BAB71493).
FT CONFLICT 189 189 R -> W (IN REF. 1; BAB71493).
FT CONFLICT 189 189
SQ SEQUENCE 274 AA; 30081 MW; ED820002E426E81A CRC64;

Query Match 13.2%; Score 70.5; DB 1; Length 274;
Best Local Similarity 33.3%; Pred. No. 6.1;
Matches 22; Conservative 2; Mismatches 15; Indels 27; Gaps 4;

FT DOMAIN 31 185 F5/8 TYPE C. 82
FT DOMAIN 377 415 GLY/PRO-RICH.
FT DOMAIN 476 601 GLY/PRO-RICH.
FT DOMAIN 610 905 PROTEIN KINASE.
FT NP BIND 616 624 ATP (BY SIMILARITY).
FT BINDING 655 655 ATP (BY SIMILARITY).

Query Match 13.3%; Score 71; DB 1; Length 913;
Best Local Similarity 29.6%; Pred. No. 18;
Matches 29; Conservative 8; Mismatches 29; Indels 32; Gaps 6;

QY 1 PGPRTRPKA-PSHAPLL-GLA-----LLRMEQRHPRARCPPLCVAGILACGFL 47
DB 483 FYQEPRGRGNPPHSPAPCVNGSALLSNPFAVRLLLATVARPRGPGPP----- 530

QY 48 LGCWG-PSHQOQCLQALEPQAVSSYLSGAPLKCRPP 84
DB 531 TPAWAKPTNTQAYSGDYMEPE-----XPGAPLLPFP 562

RESULT 11
NQOC_THETH STANDARD; PRT; 606 AA.
AC Q56227;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NADH-quinone oxidoreductase chain 12 (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain 12) (NDH-1, chain 12).
GN NQO12.
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OC NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=97172490; PubMed=9020134;
RA Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
RT thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
RT sequence of the gene cluster and thermostable properties of the
RT expressed NQO2 subunit";
RL J. Biol. Chem. 272:4201-4211 (1997).
CC -1- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain. The
CC immediate electron acceptor for the enzyme in this species is
CC believed to be menaquinone. Couples the redox reaction to proton
CC translocation (for every two electrons transferred, four hydrogen
CC ions are translocated across the cytoplasmic membrane), and thus
CC conserves the redox energy in a proton gradient.
CC -1- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNITS NQO7-14
CC CONSTITUTE THE MEMBRANE SECTOR OF THE COMPLEX.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC OF CHLOROPLASTS OR MITOCHONDRIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U52917; AAA97949.1; -
CC PIR; T11909; T11909.
CC InterPro; IPR003916; NADHub oxreds.
CC InterPro; IPR001750; Oxidored_q1.
CC InterPro; IPR001516; Oxidored_q1_N.
CC Pfam; PF00361; oxidored_q1; 1.
CC Pfam; PF00662; oxidored_q1_N; 1.
CC PRINTS; PR01434; NADHGNASES.
CC Oxidoreductase; NAD; Quinone; Transmembrane.
CC TRANSMEM 1 21 POTENTIAL.
FT TRANSMEM 29 49 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 126 146 POTENTIAL.
FT TRANSMEM 165 185 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
FT TRANSMEM 234 254 POTENTIAL.
FT TRANSMEM 266 286 POTENTIAL.


```

FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 325 345 POTENTIAL.
FT TRANSMEM 364 384 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 441 461 POTENTIAL.
FT TRANSMEM 492 512 POTENTIAL.
FT TRANSMEM 584 604 POTENTIAL.
SQ SEQUENCE 606 AA; 65141 MW; 473188P665295C CRC64;

Query Match 13.2%; Score 70.5; DB 1; Length 606;
Best Local Similarity 30.9%; Pred. No. 13;
Matches 21; Conservative 7; Mismatches 23; Indels 17; Gaps 2;

QY 15 PLGLALLRHQHRPRCPLCVAGILACGLGCGW---GPSHFQSCQLQALPQAVS 70
DB 10 PLGLFALLGFGKMRPPLCGVLASGLVLAFLGAGLLSGGARFQEWL----- 60

QY 71 SYLSPGAP 78
DB 61 ----PGIP 64

RESULT 12
AVR2_BOVIN STANDARD; PRT; 513 AA.
AC Q28043;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Activin receptor type II precursor (EC 2.7.1.37) (ACTR-II).
GN ACVR2 OR ACTRII.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Testis;
RX MEDLINE=95203477; PubMed=7534730;
RA Echier J.F., Houde A., Lussier J.G., Silversides D.W.;
RT "Bovine activin receptor type II cDNA: cloning and tissue
RT expression.";
RL Mol. Cell. Endocrinol. 106:1-8(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=9702546; PubMed=8875905;
RA Montegudo L.V., Heriz A., Flavin N., Rogers M., Ennis S.,
RA Arruga M.V.;
RT "Fluorescent in situ localization of the bovine activin receptor type
RL IIA locus on chromosome 2 (2q2.3-2.4).";
RM Mamm. Genome 7:869-869(1996);
CC -1- FUNCTION: Receptor for activin A, activin B and inhibin A.
CC -1- Involved in transmembrane signaling.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- SUBUNIT: Interacts with AIP1. Part of a complex consisting of
CC AIP1, ACVR2, ACVR1B and MADH3 (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. TGFBR
CC receptor subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; L21717; AAA74597.1; -.
CC ENBL; U43208; AAC48694.1; -.
CC PIR; I45850; I45850.

```

```

HSP; P27038; 1BTE.
DR InterPro: IPR000472; Activin receptor.
DR InterPro: IPR000333; Actn_receptorII.
DR InterPro: IPR000719; Prot_kinase.
DR InterPro: IPR008271; Ser_thr_kin_AS.
DR Pfam: PF01064; Activin_rec; 1.
DR Pfam: PF00669; pkinase; 1.
DR PRINTS: PD0653; ACTIVIN2R.
DR PRODOM: PD00001; Prot_kinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR Receptor; Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 513 ACTIVIN RECEPTOR TYPE II.
FT DOMAIN 20 135 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 136 161 POTENTIAL.
FT DOMAIN 162 513 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 192 485 PROTEIN_KINASE.
FT NP_BIND 198 206 ATP (BY SIMILARITY).
FT BINDING 219 219 ATP (BY SIMILARITY).
FT ACT_SITE 322 322 BY SIMILARITY.
FT DISULFID 30 60 BY SIMILARITY.
FT DISULFID 50 78 BY SIMILARITY.
FT DISULFID 85 104 BY SIMILARITY.
FT DISULFID 91 103 BY SIMILARITY.
FT DISULFID 105 110 BY SIMILARITY.
FT CARBOHYD 43 43 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 513 AA; 57951 MW; C2969A54CF00617B CRC64;

Query Match 13.1%; Score 70; DB 1; Length 513;
Best Local Similarity 27.7%; Pred. No. 13;
Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4;

QY 4 PTPKAPSHGAPLLGLLRHQRHPRCPLCVAGILACGLGCGWGFHFQSCQLQA 63
DB 129 PVTPKPYNYLLYSLV-----PLMLAGIVICAF-----WYRHHKWAYPPV 171

QY 64 LEFQAVSSYLSPGAPLKRPPSP 86
DB 172 LVPTQ-----DPGPP---PPSP 185

RESULT 13
AVR2_HUMAN STANDARD; PRT; 513 AA.
AC P27037; Q92474;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Activin receptor type II precursor (EC 2.7.1.37) (ACTR-II) (ACTRIIA).
GN ACVR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=92182002; PubMed=1311955;
RA Matzuk M.M., Bradley A.;
RT "Cloning of the human activin receptor cDNA reveals high evolutionary
RT conservation.";
RL Blochm. Biophys. Acta 1130:105-108(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=92231944; PubMed=1314589;
RA Donaldson C.J., Mathews L.S., Vale W.W.;
RT "Molecular cloning and binding properties of the human type II
RT activin receptor.";

```

Biochem. Biophys. Res. Commun. 184:310-316(1992).

[3] SEQUENCE FROM N.A.
RP TISSUE=Mammary gland;
RA Geiser A.G.;
RRL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
RN [4]

SEQUENCE FROM N.A.
RP Iimura T., Oida S.;
RA Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: Receptor for activin A, activin B and inhibin A.
CC Involved in transmembrane signaling.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- SUBUNIT: Interacts with AIPI1. Part of a complex consisting of
CC AIPI1, ACVR2, ACVR1B and MADH3 (By similarity).
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. TGFB
CC receptor subfamily.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

EMBL; X63128; CAA44839.1; -;
DR EMBL; X82381; CAA44245.1; -;
DR EMBL; M93415; AAA35504.1; -;
DR EMBL; D31770; BAA06548.1; -;
DR PIR; JQ1486; JQ1486.
DR HSP; P27038; IBTE.
DR Genew; HGNC:173; ACVR2.
DR MIM; 102581; -;
DR GO; GO:0005887; C:integral to plasma membrane; TAS.
DR GO; GO:0004675; F:transmembrane receptor protein serine/threo. .; TAS.
DR GO; GO:0007178; P:transmembrane receptor protein serine/threo. .; TAS.
DR InterPro; IPR000472; Activin receptor.
DR InterPro; IPR000333; Actn_receptorII.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF01054; Activin recp; 1.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00653; ACTIVIN2R.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR Receptor; Transferrase; Serine/threonine-protein kinase; ATP-binding;
KW Transmembrane; Glycoprotein; Signal
FT SIGNAL 1 19
FT CHAIN 20 513
FT DOMAIN 20 135
FT TRANSMEM 136 161
FT DOMAIN 162 513
FT DOMAIN 192 485
FT NP_BIND 198 206
FT BINDING 219 219
FT ACT_SITE 322 322
FT DISULFID 30 60
FT DISULFID 50 78
FT DISULFID 85 104
FT DISULFID 91 103
FT DISULFID 105 110
FT CARBOHYD 43 43
FT CARBOHYD 66 66
FT CONFLICT 13 13
FT CONFLICT 204 206
FT CONFLICT 348 348
FT SEQUENCE 513 AA; 57847 MW; A89822E880979618 CRC64;
SQ

Query Match 13.1%; Score 70; DB 1; Length 513;

Best Local Similarity 27.7%; Pred. No. 13;
Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4

QY 4 PTRPKAPSHSAPLLGLALLRLMRHPRARCPPLCVAGILACGLGCGWSPHFQSCSLQA 63
DB 129 PVTPEPPYNYLLYSLV-----PLMLIAGIVICAF---WVYRHHKMAVPPV 171

QY 64 LEPQAVSSYLSGAPLKGPPSP 86
DB 172 LVPTQ-----DRGPP-----PPSP 185

RESULT 14
AVR2_MOUSE
ID AVR2_MOUSE STANDARD; PRT; 513 AA.
AC P27038;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Activin receptor type II precursor (EC 2.7.1.37) (ACTR-II).
OS ACVR2 OR ACVR2A.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91256317; PubMed=1646080;
RA Mathews L.S., Vale W.W.;
RT "Expression cloning of an activin receptor, a predicted transmembrane
RL serine kinase.";
RL Cell 65:973-982(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS) OF 25-121.
RX MEDLINE=99101377; PubMed=9896286;
RA Greenwald J., Fischer W.H., Vale W.W., Choe S.;
RT "Three-finger toxin fold for the extracellular ligand-binding domain
RT of the type II activin receptor serine kinase.";
RN Nat. Struct. Biol. 6:18-22(1999).
RN [3]
RP DISULFIDE BONDS OF EXTRACELLULAR DOMAIN.
RX MEDLINE=99376271; PubMed=1049041;
RA Fischer W.H., Greenwald J., Park M., Craig A.G., Choe S., Vale W.;
RT "The disulfide bond arrangement in the extracellular domain of the
RL activin type II receptor.";
RL J. Protein Chem. 18:437-446(1999).
RN [4]
RP INTERACTION WITH AIPI1 AND IDENTIFICATION IN A COMPLEX WITH ACVR1B;
AIPI1 AND MADH3.
RX MEDLINE=20148748; PubMed=10681527;
RA Shoji H., Tsuchida K., Kishi H., Yamakawa N., Matsuzaki T., Liu Z.,
RA Nakamura T., Sugino H.;
RT "Identification and characterization of a PDZ protein that interacts
RT with activin types II receptors.";
RL J. Biol. Chem. 275:5485-5492(2000).
CC -|- FUNCTION: Receptor for activin A, activin B and inhibin A.
CC Involved in transmembrane signaling.
CC -|- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -|- SUBUNIT: Interacts with AIPI1. Part of a complex consisting of
CC AIPI1, ACVR2, ACVR1B and MADH3.
CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
CC -|- TISSUE SPECIFICITY: BRAIN, TESTIS, INTESTINE, LIVER, AND KIDNEY.
CC -|- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. TGFB
CC receptor subfamily.
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

Query Match

```

DR EMBL; M65287; AAA37171.1; .
DR PIR; A39896; A39896.
DR PDB; 1ETE; 09-FEB-99.
DR MGD; MGI:102806; Acv12.
DR STRAIN-FVB/N; TISSUE=Kidney;
DR InterPro; IPR000472; Activin_receptor.
DR InterPro; IPR000333; Actn_receptorII.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF01064; Activin_rec; 1.
DR Pfam; PF00069; Pkinase; 1.
DR PRINTS; PR00653; ACTIVIN2R.
DR ProDom; PD000001; Prot_kinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR Receptor; transferase; Serine/threonine-protein kinase; ATP-binding;
KW Transmembrane; Glycoprotein; Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 513
FT DOMAIN 20 135
FT TRANSMEM 136 161
FT DOMAIN 162 513
FT NP_BIND 192 485
FT BINDING 219 219
FT ACT_SITE 322 322
FT DISULFID 30 60
FT DISULFID 50 78
FT DISULFID 85 104
FT DISULFID 91 103
FT DISULFID 105 110
FT CARBOHYD 43 43
FT CARBOHYD 66 66
FT STRAND 29 34
FT TURN 35 36
FT HELIX 37 40
FT TURN 41 41
FT STRAND 45 49
FT STRAND 59 67
FT TURN 68 69
FT STRAND 70 79
FT STRAND 91 93
FT STRAND 101 105
FT TURN 108 109
FT HELIX 110 112
FT STRAND 114 116
FT SEQUENCE 513 AA; 57889 MW; 475CD292506BAC61 CRC64;

Query Match 13.1%; Score 70; DB 1; Length 513;
Best Local Similarity 27.7%; Pred. No. 13;
Matches 23; Conservative 7; Mismatches 27; Indels 26; Gaps 4;

QY 4 PTPKAPSHSAPLLGLALLRHQHPRCPCPLCVAGILACGFLGCGWGFHFQOQSLQA 63
DB 129 FVTPKPYNYLLVSLV-----PMLIAGIVICAF----WYRHHKQWAPPV 171
QY 64 LEPQAVSSYLPAGPLKGRPPSP 86
DB 172 LVPTQ-----DFGPP-----PPSP 185

RESULT 15
AT15_MOUSE
ID AT15_MOUSE STANDARD; PRT; 562 AA.
AC P59384;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
DE thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).
GN ADAMTS15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; BC043308; AAH43308.1; .
DR MGD; MGI:2449569; Adamts15.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR01762; Disintegrin.
DR InterPro; IPR01818; Pept_M12B_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS00215; ADAM_MEPRO; PARTIAL.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Repeat;
KW Extracellular matrix.
FT NON_TER 1
FT DOMAIN <1 39
FT DOMAIN 40 127
FT DOMAIN 128 183
FT DOMAIN 184 312
FT DOMAIN 311 450

```

```
FT DOMAIN 451 507 TSP TYPE-1 2.
FT DOMAIN 508 561 TSP TYPE-1 3.
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 562 AA; 60963 MW; 48009AD37D4E4A CRC64;

Query Match 13.1%; Score 70; DB 1; Length 562;
Best Local Similarity 32.9%; Pred. No. 14;
Matches 25; Conservative 9; Mismatches 28; Indels 14; Gaps 4;

QY 3 RPTRPKAPSHSAPLLGLALLRMHOR--HPRARCPLCVAGILACGFLIGCWGPSHFQOSC 60
Db : ||||| : || : : : : |||||
420 KSTRPKDP-RGSPVLRNSVLNSNQVEQPDNRPPARWVA-----GSWGPC--SVSC 467

QY 61 LQALEPQAYSSYLSPG 76
Db : : |||
468 GSGLQKRAVDCRDSFG 483
```

Search completed: March 13, 2004, 07:39:56
Job time : 4.02083 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 9.59896 Seconds
(without alignments)
3188.394 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534

Sequence: 1 PGRTPRPKAPSHSAPLLGLA.....PLKGRPPSPGFORQRRQR 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	414	77.5	1427	4 Q96L37	Q96L37 homo sapien
2	95.5	17.9	225	4 Q8BYE0	Q8BYE0 homo sapien
3	85.5	16.0	225	11 Q8BK72	Q8BK72 mus musculus
4	84.5	15.8	225	11 Q99JA6	Q99JA6 mus musculus
5	82.5	15.4	281	10 Q8H468	Q8H468 oryza sativ
6	82.5	15.4	1187	10 Q9LWU2	Q9LWU2 oryza sativ
7	81.5	15.3	649	10 Q7XDV0	Q7XDV0 oryza sativ
8	81	15.2	259	4 Q96D16	Q96D16 homo sapien
9	81	15.2	1071	4 Q9H407	Q9H407 homo sapien
10	79.5	14.9	249	11 Q8C8J4	Q8C8J4 mus musculus
11	79.5	14.9	341	11 Q8BG91	Q8BG91 mus musculus
12	78.5	14.7	186	15 Q7W980	Q7W980 bordetella
13	77.5	14.5	210	13 Q73592	Q73592 gallus gall
14	77.5	14.5	598	16 Q8VKV7	Q8VKV7 mycobacteri
15	77.5	14.5	992	11 Q8BNJ0	Q8BNJ0 mus musculus
16	76.5	14.3	169	6 Q9TUH6	Q9TUH6 hylobates l

Q7vbk0 bordetella
Q9tuH5 hylobates p
Q8lq19 oryza sativ
Q64108 ratus sp.
Q9vzw7 drosophila
Q8ign3 drosophila
Q8u65 mus musculu
Q75851 homo sapien
Q8n119 homo sapien
Q7fin8 rhesus cyto
Q8yu84 anabaena sp
Q8n420 homo sapien
Q86v22 homo sapien
Q93s18 myxococcus
Q88ve0 mus musculu
Q9h8w7 homo sapien
Q9bvg7 homo sapien
Q9tuH8 hylobates p
Q9tuH7 hylobates p
Q82e66 streptomyce
Q8466 mus musculu
Q8xxx4 homo sapien
Q8nce1 homo sapien
Q8xtj8 raistonia s
Q9lmq6 arabidopsis
Q8t8n5 hylobates s
Q8n9p6 homo sapien
Q89r45 bradyrhizob
Q9aeh9 chlorobium

ALIGNMENTS

RESULT 1

Q96L37 ID Q96L37 PRELIMINARY; PRT; 1427 AA.

AC Q96L37; DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,
Fujikawa K.,
RT "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a
RT Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura.", a
RL J. Biol. Chem. 276:41059-41063(2001).
DR EMBL; AY055376; AAU17652.1; -.
DR MEROPS; M12.241; -.
DR Genew; HGNC:1366; ADAMTS13.
DR GO; GO:004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSF1.
DR Pfam; PF01421; Reprolysin; 1.
DR SMART; SM00209; TSP1; 7.
DR PROSITE; PS50215; ADAM_MERO; 1.
DR PROSITE; PS50092; TSP1; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Signal.
FT SIGNAL 1 33 POTENTIAL.

```

FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB1BC3AABCLIA4442 CRC64;

Query Match
Best Local Similarity 77.5%; Score 414; DB 4; Length 1427;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MHQHRPRAPCPPLCVAGILACGFLGCGWPSHFQOCLQALEPQAVSSYLSGAPLKGRP 83
DB 1 MHQHRPRAPCPPLCVAGILACGFLGCGWPSHFQOCLQALEPQAVSSYLSGAPLKGRP 60

QY 84 PSPGFQQRQRR 97
DB 61 PSPGFQQRQRR 74

RESULT 2
Q9BYEO PRELIMINARY; PRT; 225 AA.
AC Q9BYEO;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BHLH factor Hes7.
GN HES7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21160566; PubMed=11260262;
RA Bessho Y., Miyoshi G., Sakata R., Kageyama R.;
RT "Hes7: a bHLH-type repressor gene regulated by Notch and expressed in
the presomitic mesoderm.";
RL Genes Cells 6:175-185(2001).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
DR EMBL; AB049064; BAB39525.1; -.
DR Genbank; HGNC:15977; HES7.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0003677; F:DNA binding; NAS.
DR GO; GO:0007498; P:mesoderm development; NAS.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00886; HLH_2; 1.
SQ SEQUENCE 225 AA; 24894 MW; 027B5BAAB4AC8247 CRC64;

Query Match
Best Local Similarity 17.9%; Score 95.5; DB 4; Length 225;
Matches 34; Conservative 5; Mismatches 35; Indels 35; Gaps 5;

QY 1 PGRP-----TRPKAPGHS-----APLGLALLR---MHQHRPRAPCPPLCVAGILACGFLG 49
DB 130 PPRPKVDPDRPPAPRSLDPAAPALGPALHQRPPVHQGHSPRC-----173

QY 50 CWGSHFQOCLQALEPQAVSSYLSP-----GAPLKGRPPSPGFQR 90
DB 174 AWSFSLCSFRAAGSAPAPLTLGLPPPPHQRQDGAPKAPLPPPPAFWR 222

RESULT 3
Q8BKT2 PRELIMINARY; PRT; 225 AA.
AC Q8BKT2;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BHLH factor Hes7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK050791; BAC34412.1; -.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00886; HLH_2; 1.
SQ SEQUENCE 225 AA; 24889 MW; 4B22E7D71A7FF5E8 CRC64;

Query Match
Best Local Similarity 16.0%; Score 85.5; DB 11; Length 225;
Matches 33; Conservative 6; Mismatches 37; Indels 23; Gaps 6;

QY 1 PGRPTRPKAP-SHSAPLLGLALLRMHQRHPRAPCPPLCVAGILACGFLGCGWPSHFQO 59
DB 138 PGLPA-PRPPLDPASPILGPAL---HQRPPVHQGPP---SPRLA-----WSPSHCSSR 183

QY 60 C-----LQALEPQAVSSYLSGAPLKGRPPSPGFQR 90
DB 184 AGDSGAPAPLTLGLPPPPPYRQDGAPKAPLPPPPAFWR 222

RESULT 4
Q99JAG PRELIMINARY; PRT; 225 AA.
AC Q99JAG;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BHLH factor Hes7.
GN HES7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bessho Y., Miyoshi G., Kageyama R.;
RT "Characterization of novel BHLH genes.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21160566; PubMed=11260262;
RA Bessho Y., Miyoshi G., Sakata R., Kageyama R.;
RT "Hes7: a bHLH-type repressor gene regulated by Notch and expressed in
the presomitic mesoderm.";
RL Genes Cells 6:175-185(2001).
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
TRANSCRIPTION FACTORS.
DR EMBL; AB050104; BAB41134.1; -.
DR EMBL; AB049065; BAB39526.1; -.
DR MGD; MGI:2135679; Hes7.
DR GO; GO:0016564; F:transcriptional repressor activity; IDA.
DR GO; GO:0007219; P:N signaling pathway; IDA.
DR GO; GO:0000122; P:negative regulation of transcription from P...; IDA.
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00038; HLH_1; 1.
DR PROSITE; PS00886; HLH_2; 1.
SQ SEQUENCE 225 AA; 24899 MW; 4B22E7D4D57FF5E8 CRC64;

Query Match
Best Local Similarity 15.8%; Score 84.5; DB 11; Length 225;

```

Best Local Similarity 33.3%; Pred. No. 0.17;
Matches 33; Conservative 6; Mismatches 37; Indels 23; Gaps 6;

1 PGRPRKAP-SHSAPLLGLLRMHQHRARCPCLVAGILACGFLLGWGPSSHQQS 59
||| ||| :|||
138 FGLPA-PFPLDPASFLIGPAL---HQRPVHQPP-----SPRLA-----WSFSHCSSR 183

60 C-----LQALEPQAVSSYLSPGAPLKGRPPPSGFQR 90

184 AGDGGAPAPLTGLLPFPFPFPVRQDGAKPAPLEPPAPFR 222

```

RESULT 5
ID      Q8H468
NAME    Q8H468
PRT;    281 AA.
PRELIMINARY;
Q8H468;
01-WAR-2003 (TREMELrel. 23, Created)
01-WAR-2003 (TREMELrel. 23, Last sequence update)
01-WAR-2003 (TREMELrel. 23, Last annotation update)
P0470D12.26 protein.
P0470D12.26.
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
NCBI taxID=39947;
[1]
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, PAC
clone:P0470D12.";
Submitted (OCT-2001) to the ENBL/GenBank/DBJ databases.
EMBL; AF0904300; BAC15992.1; -.
SEQUENCE 281 AA; 30264 MW; 3BED63828FFDF973E CRC64;

```

```
Query Match      15.4%; Score 82.5; DB 10; Length 281;  
Best Local Similarity 28.6%; Pred. NO. 0.36;  
Matches 36; Conservative 7; Mismatches 39; Indels 43; Gaps 6;
```

1 PGRTRPKAP-----SHSAPLGLALLRRHQHPRARCP--PLCVAGIIACGFLIGCW 51
 | | |
31 PNLPORRAARIAAAASISHHTSALGLELRTVARWPAVPSRRLAAPLLA----- 82
 | : |
52 GPSHFQQSCLOALEPQAVSYSLSP-----GAPLKGRPSPSGFOR-----QR 92
 | : |
83 --GH-----RRAAAPRTAIHTSPATARSWRPTSSCPPRAAPSPSFVRTPPAMESQE 135

93 QRQR 97
 : | |
136 RRNR 140

RESULT 6
Q9LWU2 PRELIMINARY; PRT; 1187 AA.
Q9LWU2
01-OCT-2000 (TRENBLrel. 15, Created)
01-OCT-2000 (TRENBLrel. 15, Last sequence update)
01-OCT-2002 (TRENBLrel. 22, Last annotation update)
Hypothetical protein.
Oryza sativa (Rice).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Eriarctoidae; Oryzeae; Oryza.
NCBI_TaxID=4530;
{1}
SEQUENCE FROM N.A.
STRAIN=cv. Nipponbare;
Sasaki T., Matsumoto T., Yamamoto K.;
"Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, PAC
clone:R0541H01."
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

```

DR EMBL; AP001389; BAA92718.1; -.
DR Gramene; Q9LWU2; -.
DR InterPro; IPR005213; HGWP.
DR Pfam; PF03578; HGWP; 9.
DR Hypothetical protein.
SQ SEQUENCE 1187 AA;  EDC8A0AA47CG920 CRC64;
      15.4%; Score 82.5; DB 10; Length 1187;
      Best Local Similarity 29.5%; Pred. No. 1.5;
      Matches 36; Conservative 11; Mismatches 34; Indels 41; Gaps 8
Qy      3 RPRPKAPSHSAPLGLALLRMQRRHPA-----RC-----PPLCV--- 38
Db      309 RPVFPFARLAYAAGV-GHIFRFRQPSPLTGLACTASILCRRCWASTSSPPACLRL 367
Qy      39 -AGILAC--GFLL-GCWGSPSHFOOSCL---QALEPQAVSY-----LSPGAPLKGR 82
Db      368 PTGVLTACTALLCRRCWASTSSRPACLCRRRPVFPFAPLAYAAVTRCLLSPPACLRR 427
Qy      83 PP 84
Db      428 RP 429

```

```

RESULT 7
Q7XDVO
ID ID Q7XDVO PRELIMINARY; PRT; 649 AA.
AC Q7XDVO;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN OSUNBAR0068N06.22.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoidae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
  chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR ENBL; AE017100; AAP54041.1; -.
KW Hypothetical protein.
SQ SEQUENCE 649 AA; 71149 MW; EEB232805DC0D778 CRC64;

```

```
Query Match          15.3%; Score 81.5; DB 10; Length 649;
Best Local Similarity 28.8%; Pred. No. 1.1;
Matches 30; Conservative 12; Mismatches 33; Indels 29; Gaps 6;
```

Qy	2	GRTRPKAPSHSA-----PILGALLRMHORHP-RARCPPLCVAGILA--C 44
		: : : : :
Dd	230	GRPLHRRPASVAAEKCLRHLHGWSINPLLGV-----HDTAGLPLPPSGASAYTA 281
		: : : : : : : : : : : : : : :
Qy	45	GFLL-CCWFSPHFQQCSIQALEPOAVSSYLSPGAKLGRRPPSG 87
		: : : : : : : : : : : : : : :
Dd	282	GLSCRCRWASTTTPPALCRRCRVWTSQNTITGSP---RPPPTG 322

RESULT 8	
Q96D16	
ID	Q96D16
AC	PRELIMINARY; PRT; 259 AA.
Q96D16	
DT	01-DEC-2001 (T=EMBLrel. 19, Created)
DT	01-DEC-2001 (T=EMBLrel. 19, Last sequence update)

```

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Similar to hypothetical protein FLJ11467.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAUSBERG R.;
RA Strausberg R.;
DR EMBL: BC013435; AALH1435.1; -.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 26945 MW; F954FCEA861CC8C7 CRC64;

Query Match 15.2%; Score 81; DB 4; Length 259;
Best Local Similarity 32.5%; Pred. No. 0.49;
Matches 37; Conservative 6; Mismatches 25; Indels 46; Gaps 8;

QY 3 RPTREPKA---PSHAPILGLALLRMHQHPRARCPPLCVAGILACGFLIGC-----WG 52
DB 163 RPLVPGAEPPVLAAPVGLS-----QRHPPARCR---AG-LHGSLLPAGCHVPPVHRGV 212
QY 53 PSHFQSCLOALEPQAVSSYLSPGAPLK-----GRPP-SPG 87
DB 213 PRHLQ-----EPAVASPOLGGAARVKKRHLPSPAPRGRPRHPPGAPG 259

RESULT 9
Q9H407 PRELIMINARY; PRT; 1071 AA.
AC Q9H407;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE D1492J12.1 (Novel protein similar to zinc finger protein human
DE immunodeficiency virus type I enhancer-binding protein 1 (HIVEP1))
DE (Fragment).
GN D1492J12.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Peck A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL121919; CAC12920.1; -.
DR HSP; P15822; 1BBO.
DR Genew; HGNC:16167; C20orf174.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2_2.
DR SMART; SM00355; Znf_C2H2_2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
DR Metal-binding; Zinc; Zinc-finger.
FT NON_TER 1
FT NON_TER 1071
SQ SEQUENCE 1071 AA; 113759 MW; 470EC64E39C4EABC CRC64;

Query Match 15.2%; Score 81; DB 4; Length 1071;
Best Local Similarity 31.6%; Pred. No. 1.9;
Matches 31; Conservative 10; Mismatches 31; Indels 26; Gaps 6;

QY 1 PGRTRPKAPSHSAPLLGLA-----LLRMHQHPRARCP-PLCVAGILACGFLGCGWGPS 54
DB 782 PSLATPPQAPRVLSALADNAPSKVLLRL---PQAETPLPLIP-----WGPR 826
QY 55 HFQSCLOALEPQAVSSYLSPP--GAPLKGRP---PSPG 87
DB 827 HSQDSLCSGWFPERASFVGSGLGTLPSPASGSPG 864
```

```

RESULT 10
Q8C8J4 PRELIMINARY; PRT; 249 AA.
ID Q8C8J4;
AC Q8C8J4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc finger protein of the cerebellum 4.
GN ZIC4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK046913; BAC32918.1; -.
DR MGD; MGI:107201; Zic4.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 2.
DR PRODOM; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 1.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
SQ SEQUENCE 249 AA; 27513 MW; 0B705BF6BA8FA9AA CRC64;

Query Match 14.9%; Score 79.5; DB 11; Length 249;
Best Local Similarity 27.9%; Pred. No. 0.69;
Matches 31; Conservative 12; Mismatches 49; Indels 19; Gaps 4;

QY 1 PGRTRPKAPSHSAPLLGLALL-----RMHQH-----PRACPLCVAGILACGF---- 46
DB 44 PGLHEQPPQASHSRPLNGLLGLPGDWYARSEFPAPGPMARSDTLATALLH-GYGMN 102
QY 47 ----LLCGWSPHFQSCLOALEPQAVSSYLSPGAPLKGRPPSPGFQRQRQ 93
DB 103 LTWNLTAPHGPGAFRYRQPIKQELICKLWGLDDSPRPSCKTFSTWHE 153

RESULT 11
Q8BG91 PRELIMINARY; PRT; 341 AA.
ID Q8BG91;
AC Q8BG91;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc finger protein of the cerebellum 4.
GN ZIC4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028305; BAC25871.1; -.
DR EMBL; AK035908; BAC29239.1; -.
DR MGD; MGI:107201; Zic4.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 4.
DR PRODOM; PD000003; Znf_C2H2; 1.
```



```
DR SMART; SM00355; ZnF C2H2; 5.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 4.
SQ SEQUENCE -341 AA; 37337 MW; B4B510DE736230CF CRC64;

Query Match
Best Local Similarity 14.9%; Score 79.5; DB 11; Length 341;
Matches 31; Conservative 12; Mismatches 49; Indels 19; Gaps 4;

QY 1 PGRTTRKAPSHSAPLLGLALL-----RMQHRH-----PRACPLCVAGILACGPF---- 46
DB 51 PGLHEQPQASHSRPLNGLLGLGIPGDMYARSEFPAPGFMARSDTLATATLH-GYGMN 109
QY 47 ----LLCGWSPHFQOSCLQALEPQAVSSYLSPGAPLKGRRPSPGFORQ 93
DB 110 LTWNLTAPHGCAFRYVRQPIKQLICKWLGDSPMRPCSKPTSTWHE 160

RESULT 12
QY 07W980 PRELIMINARY; PRT; 186 AA.
AC Q7W980;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE L-2,4-diaminobutyric acid acetyltransferase (EC 2.3.1.-).
CN ECTA OR BPP1888.
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271;
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Cerdeno-Farraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Chervach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabbintwitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica."
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640428; CAE37189.1; -.
KW Acyltransferase; Transferase; Complete proteome.
SQ SEQUENCE 186 AA; 20729 MW; C0B7C716917F9CF4 CRC64;

Query Match
Best Local Similarity 14.7%; Score 78.5; DB 16; Length 186;
Matches 28; Conservative 12; Mismatches 28; Indels 25; Gaps 5;

QY 6 RPKAPSHSAPLLGLA-----LLRMQHRH-----ACPLCVAGILACGFLGC 50
DB 2 RKQETSTSPDISVQAPASALRYHURPFRNDCAIHQLVSECPDLNSLYA--YLLLC 59
QY 51 WGPSPHFQOSCLQALEPQA-----VSSYLSPGAP 78
DB 60 ---EHAHTCVVAESPGRGIDGFVSAYLLPTRP 89

RESULT 13
QY 073592 PRELIMINARY; PRT; 210 AA.
AC Q73592;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CMIX.
```

```
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RA Peale F.V., Sugden L., Bothwell M.;
RT "Characterization of CMIX, a chicken homeobox gene related to the
RT Xenopus gene Mix.1.";
RL Mech. Dev. 75:179-182(1998).
RN [2]
SQ SEQUENCE FROM N.A.
RC STRAIN=White leghorn;
RX MEDLINE=98141813; PubMed=9473273;
RA Peale F.V. Jr., Mason K., Hunter A.W., Bothwell M.;
RT "Multiplex display polymerase chain reaction amplifies and resolves
RT related sequences sharing a single moderately conserved domain.";
RL Anal. Biochem. 256:158-168(1998).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; U34615; AAC36453.1; -.
DR HSSP; P06601; 1FJL.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR007104; Paired homeo.
DR Pfam; PF00046; homeobox; 1.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT VARIANT 154 154 P->S.
SQ SEQUENCE 210 AA; 21981 MW; F8A4617E2385B5A9 CRC64;

Query Match
Best Local Similarity 14.5%; Score 77.5; DB 13; Length 210;
Matches 26; Conservative 2; Mismatches 31; Indels 33; Gaps 3;

QY 1 PGRTTRKAPSHSAPLLGLALLRMQHRH-----PRACPLCVAGILACGFLCGWSPSHF 56
DB 122 PGAPAPPPPPQSPQSCGAAPLLRAFEHREHNPRAAGP----- 160
QY 57 QQSCQLQALEPQAVSSYLSPGAPLKGRRPSPGF 88
DB 161 ----GSALRPHGSG-----GAPAGYPYPPAPF 184

RESULT 14
QY 08VKN7 PRELIMINARY; PRT; 598 AA.
AC Q8VKN7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein MT0291.4.
GN MT0291.4.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
SQ SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
```

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE006937; AAK44515.1; -.
 DR TIGR; MT0291; -.
 KW Hypothetical protein.
 SQ SEQUENCE 598 AA; 58523 MW; 75293DE1933B0ED6 CR664;

Query Match	14.5%;	Score	77.5;	DB	16;	Length	598;
Best local Similarity	28.8%;	Pred.No.	2.7;				
Matches	38;	Conservative	7;	Mismatches	40;	Indels	47; Gaps 7;
Qy	1	PGRTRTKAP-----SH-----SAPLLG--LALLRMHQHPRARCPCPLCVAGILAC	44				
Db	163	PSPPSAFVAPWDPDLPPLNNHPAPPAPVEGVFLAELPISGRPVRAWGSLIARICCC	222				
Qy	45	----GFLIGCWGSPGHQQSC-----LQALEPQAVSSYLSFGAPL----	79				
Db	223	RVCSGVLGALNPGR-PSSCPKPPAPAVPAGVPVPLPPLPISTELFPAPPLPLP	281				
Qy	80	-----KORPSP	86				
Db	282	ALPTSPGAPPAP	293				

RESULT 15

Q8BNLO	PRELIMINARY;	PRT;	992 AA.
ID	Q8BNLO		
AC	Q8BNLO;		
DT	01-MAR-2003	(TREMBLrel. 23, Created)	
DT	01-MAR-2003	(TREMBLrel. 23, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Hypothetical Gram-positive cocci surface protein 'anchoring'		
DE	hexapeptide/RA domain/PDZ domain.		
GN	D93C005D10RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_taxid=10090;		
[1]	SEQUENCE FROM N.A.		
RP	STRAIN=C57BL/6J; TISSUE=Body;		
RC	MEDLINE=22354583; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	"Analysis of the mouse transcriptome based on functional annotation of		
RT	Nature 420:563-573 (2002).";		
RL	EMBL; AK083410; BAC38904.1; -.		
DR	PIR; PT0721; PT0721.		
DR	MGD; MGI:3443088; D93C005D10RIK.		
DR	GO; GO:0007218; P:neuropeptide signaling pathway; IEA.		
DR	InterPro; IPR002170; DIL.		
DR	InterPro; IPR000159; RA domain.		
DR	InterPro; IPR008994; SMAD_FHA.		
DR	Pfam; PF01843; DIL; 1.		
DR	Pfam; PF00788; RA; 1.		
DR	SMART; SM00314; RA; 1.		
DR	PROSITE; PS0200; RA; 1.		
DR	Hypothetical protein.		
SW	SEQUENCE	992 AA, 108964 MW; 13AB2678AB9BE1FC CRC64;	

	Query Match	14.5%	Score 77.5	DB 11	Length 992
	Best Local Similarity	27.2%	Pred. No. 4.4		
	Matches 37	Conservative 6	Mismatches 30	Indels 63	Gaps 7
QY	1	PGPRTPKAPSHSNPLICLALLRMHQHPRARCPP	-----LCVAGILACGFLG	50	
Db	878	PGRVAVETAPHSLSLVTA	-----PRAQPPQRPTQGRRSQAGSL	920	
QY	51	WGSPSHFQOQSL	-----QALEPQAVS	-----SYLS	-----PGAPLKG
Db	921	-----HTDSSCMLTPTPLGLEPAGSPWFPSGLCGRAVL	DGORNPGGGLP	GALEGD	AI 976
QY	82	---RPSPSPGFRQQR	94		

Db 977 QDAEPPAEGLOPQLQR 992

Search completed: March 13, 2004, 07:43:19
Job time : 12.599 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 14.1964 Seconds
(without alignments)
1930.574 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534

Sequence: 1 PGRPRPKAPSHAPLGLA.....PLKGRPPSPGFRQQRQR 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1990s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description.
1	534	100.0	1416	5	AB04153 Human ADA
2	492	92.1	1120	4	AAG63829 Amino aci
3	492	92.1	1120	6	ABG74113 Human mat
4	414	77.5	1427	5	AAG24450 Human Von
5	406	76.0	1427	7	ADD94038 Human agg
6	204	38.2	242	5	AAU97641 Human agg
7	204	38.2	242	7	ADD94031 Human agg
8	99.5	18.6	933	5	AAU79217 Human ADA
9	99.5	18.6	1445	5	ABG98125 Human PMW
10	93	17.4	115	5	AAC06459 Human pol
11	86.5	16.2	938	6	ABP98888 Human mol
12	86.5	16.2	1047	5	ABG31634 Human PMP
13	81	15.2	539	7	ADC32956 Human nov
14	80.5	15.1	19938	6	ABP76680 Streptomy
15	80	15.0	202	4	AAC01892 Human pol
16	80	15.0	307	4	AAU40850 Human pol
17	79	14.8	117	3	AB16741 Bacteriop
18	79	14.8	412	5	ABP42007 Human ova
19	78.5	14.7	146	5	ABP42439 Human ova
20	78	14.6	328	4	AAU80248 Human pro
21	77.5	14.5	598	6	ABU37030 Protein e
22	77	14.4	451	4	AAG73931 Human col
23	77	14.4	996	4	ABG09473 Novel hum
24	77	14.4	996	4	ABG27805 Novel hum
25	76.5	14.3	211	4	ABG19617 Novel hum

26	76.5	14.3	211	7	ADC33082	Human nov
27	76	14.2	216	5	AAU83219	Novel sec
28	76	14.2	19938	6	ABG98398	Streptomy
29	75.5	14.1	305	4	ABG27041	Novel hum
30	75	14.0	287	7	ADG40108	Human NOV
31	75	14.0	924	3	AAU52991	Drosophil
32	75	14.0	958	4	ABG62764	Drosophil
33	75	14.0	1152	5	ABG78295	Amino aci
34	74.5	14.0	165	7	ADG08117	Novel pro
35	74.5	14.0	175	4	ABG03211	Human mus
36	74.5	14.0	175	4	AAU42371	Human pol
37	74.5	14.0	175	6	ABU12505	Novel hum
38	74	13.9	236	2	AAU21834	Sequence
39	74	13.9	294	5	AAU91148	Human sec
40	74	13.9	910	7	ADD47035	Rat Prote
41	74	13.9	935	4	ABG62742	Drosophil
42	74	13.9	4123	7	ABU62079	Human jel
43	73.5	13.8	274	4	AAU39109	Human pol
44	73.5	13.8	289	4	AAU40895	Human pol
45	73.5	13.8	501	5	ABG70171	Human pre

ALIGNMENTS

RESULT 1

AB04153 ID AB04153 standard; protein; 1416 AA.

XX AC AB04153;

XX AC

DT 26-MAR-2002 (first entry)

XX AC

DE Human ADAMTS-M polypeptide.

XX AC

OS Osteoarthritis; rheumatoid arthritis; inflammatory bowel disease;
XX Homo sapiens.
FH Key
FT Peptide
FT Location/Qualifiers
FT 1..97
FT /label= Prodomain
FT /note= "The mature form of the ADAMTS-M protein is
FT processed by furin cleavage of the prodomain"
FT 94..97
FT Cleavage-site
FT /label= Furin_cleavage_site
FT 98..1416
FT Protein
FT /label= Mature ADAMTS-M protein
FT /note= "The mature form of the ADAMTS-M protein is
FT processed by furin cleavage of the prodomain"
FT 98..311
FT Domain
FT /label= Metalloprotease_domain
FT 247..272
FT Domain
FT /label= Zinc-binding_motif
FT 324..394
FT Domain
FT /label= Disintegrin_domain
FT 410..473
FT Domain
FT /label= Thrombospondin_submotif
FT 419..424
FT Domain
FT /label= Heparin-binding_domain
FT 1099..1156
FT Domain
FT /label= Thrombospondin_submotif

EP1152055-A1.

07-NOV-2001.

XX 24-APR-2001; 2001EP-00303706.
XX 27-APR-2000; 2000US-0200040P.
XX (PFIZ) PFIZER PROD INC.
XX Buckbinder L, Mitchell PG, Wachtmann TS, Walsh RT;
XX WPI; 2002-084275/12.
XX N-PSDB; ABA02549.
XX New polynucleotide, useful in gene therapy, particularly for treating or
XX preventing e.g. arthritis, Crohn's disease, Alzheimer's disease and organ
XX transplant toxicity and rejection, comprises ADAMTS polynucleotide and
XX encoded polypeptide.
XX Claim 4; Fig 2; 31pp; English.
XX The present sequence represents a ADAMTS protein, designated ADAMTS-M,
XX that exhibits the characteristics of the ADAM (A disintegrin and
XX Metalloprotease) family of metalloproteases, and contains a
XX thrombospondin domain (TS). The protein is encoded by the cDNA given in
XX ABA02549. The specification describes a newly isolated polynucleotide,
XX comprising a nucleotide sequence encoding an ADAMTS-M polypeptide as
XX given in the specification, or a metalloproteinase, disintegrin domain,
XX prodomain or its thrombospondin submotif. The polynucleotide, polypeptide
XX and agent are useful for manufacturing a medicament for treating a
XX subject in need of altering activity or expression of ADAMTS-M. The
XX polynucleotide, ADAMTS-M polypeptide and agent are useful for
XX manufacturing a medicament for treating arthritis (osteoarthritis and
XX rheumatoid arthritis), inflammatory bowel disease, Crohn's disease,
XX asthma, Alzheimer's disease, organ transplant toxicity and rejection,
XX cachexia, allergy, cancer (e.g. solid tumour cancer including colon,
XX breast, lung, prostate, brain or haematopoietic malignancies including
XX leukaemia and lymphoma), osteoporosis, atherosclerosis, aortic aneurysm,
XX congestive heart failure, myocardial infarction, stroke, head trauma,
XX spinal cord injury, neurodegenerative diseases, autoimmune disorders,
XX Huntington's disease, Parkinson's disease, migraines, pain, depression,
XX multiple sclerosis, abnormal wound healing, burns, infertility or
XX diabetic shock. The polynucleotide and polypeptide are also useful for
XX diagnosing the diseases above. The polynucleotide is particularly useful
XX in gene therapy for treating the diseases cited above
XX
XX Sequence 1416 AA;
Query Match 100.0%; Score 534; DB 5; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1.6e-46;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PGRPTRKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGFLLCGWSGSHFQQSC 60
Db 1 PGRPTRKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGFLLCGWSGSHFQQSC 60
QY 61 LQALEPAVSSYLSFGAPLKGRRPSPGQRRQRORR 97
Db 61 LQALEPAVSSYLSFGAPLKGRRPSPGQRRQRORR 97
RESULT 2
ID AAG63829
AC AAG63829 standard; protein; 1120 AA.
XX AAG63829;
XX 29-OCT-2001 (first entry)
XX Amino acid sequence of a human zdint5 polypeptide.
XX Human; zdint5; anti-angiogenic; intestinal polypeptide; wound healing;
XX extracellular matrix interaction; tumor suppression; gamete maturation;
XX immunologic recognition; gastrointestinal irradiation; chemotherapy;
XX proteolysis; apoptosis; angiogenesis; infection; cell adhesion;

KW cell fusion; cell signalling; tumour; Crohn's disease; melanoma; trauma;
KW inflammatory bowel disease; food poisoning; degenerative disease;
XX inflammation; fertility; gamete maturation; epithelial disorder.
XX Homo sapiens.
XX Key Location/Qualifiers
FH Modified-site 172 /note= "potential N-linked glycosylation site"
FT Modified-site 176 /note= "potential N-linked glycosylation site"
FT Misc-difference 474 /note= "unspecified residue encoded by TTN"
FT Modified-site 485 /note= "potential N-linked glycosylation site"
FT Modified-site 533 /note= "potential N-linked glycosylation site"
FT Modified-site 560 /note= "potential N-linked glycosylation site"
FT Modified-site 595 /note= "potential N-linked glycosylation site"
FT Modified-site 635 /note= "potential N-linked glycosylation site"
XX WO200159112-A1.
XX 16-AUG-2001.
XX 09-FEB-2001; 2001WO-US0004198.
XX 10-FEB-2000; 2000US-00501806.
XX (ZYMO) ZYMOGENETICS INC.
XX Holloway JL, Sheppard PO;
XX WPI; 2001-522477/57.
XX N-PSDB; AAH74765.
XX New anti-angiogenic intestinal polypeptides, zdint5 polypeptides, which
XX are members of disintegrin proteases, for modulating extracellular matrix
XX interaction, tumor suppression and wound healing.
XX Claim 2; Page 84-88; 92pp; English.
XX The present sequence represents a human zdint5 polypeptide. The zdint5
XX polypeptide is an anti-angiogenic intestinal polypeptide. Zdint5 is used
XX for modulating extracellular matrix interactions. Zdint5 polypeptide is
XX useful as a tool for identifying new family members of polypeptides.
XX Zdint5 polynucleotides are useful as probes or primers to clone 5' non-
XX coding regions of zdint5 gene. Zdint5 polypeptides are used for tumour
XX suppression, gamete maturation, immunologic recognition, and growth and
XX differentiation either working in isolation or in conjunction with other
XX molecules in colon, small intestine, fetal lung, testis and B-cells.
XX zdint5 polypeptides are also useful for promoting wound healing, in the
XX treatment of disorders associated with recovery after gastrointestinal
XX irradiation, chemotherapy or antibody use, as anti-infectives, and
XX extracellular matrix repair and remodeling. The polypeptides are also
XX useful for modulating proteolysis, apoptosis, angiogenesis, infection,
XX cell adhesion, cell fusion and signalling. The polypeptides are also
XX useful for treating tumour formation, Crohn's disease, inflammatory bowel
XX disease, food poisoning, melanoma, degenerative diseases, disorders
XX related to immunity, inflammation, fertility, gamete maturation,
XX immunology, trauma and epithelial disorders
XX
XX Sequence 1120 AA;
Query Match 92.1%; Score 492; DB 4; Length 1120;
Best Local Similarity 100.0%; Pred. No. 2.9e-42;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 KAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGFLLCGWSGSHFQQSCLOALEPQ 67
|||||

Db 15 KAPSHSAPLLGLALLRMHQHPRARCPPLCVAGILACGFLGCGWGPSHFQOSCLQALEPQ 74

Qy 68 AVSSYLSPGAPLKGRRPPSPGFQQRQRR 97

Db 75 AVSSYLSPGAPLKGRRPPSPGFQQRQRR 104

RESULT 3

ID ABG74113 standard; protein; 1120 AA.

XX ABG74113;

AC ABG74113;

XX 01-MAY-2003 (first entry)

DT Human mature zidnt5.

DE Human; zidnt5; ADAM; a disintegrin and metalloprotease; detergent;

XX host defence; tumour; extracellular matrix repair; proteolysis;

KW apoptosis; angiogenesis; bacterial infection; Crohn's disease;

KW inflammatory bowel disease; food poisoning; melanoma;

KW degenerative disease; chromosome 9q34; wound healing; TSP-1;

KW Thrombospondin domain.

XX Homo sapiens.

XX Key Location/Qualifiers

FH Misc-difference 474 /label= UNKNOWN

FT /note= "Encoded by TTN"

FT Misc-difference 1118..1119

FT /note= "Encoded by CTGTGAAT"

XX US2002142439-A1.

PN 03-OCT-2002.

XX 09-FEB-2001; 2001US-00781080.

XX 10-FEB-2000; 2000US-0191511P.

XX (HOLL/) HOLLOWAY J L.

PA (SHEP/) SHEPPARD P O.

PA (YAMA/) YAMAMOTO G.

XX Holloway JL, Sheppard PO, Yamamoto G;

XX WPI: 2003-174136/17.

DR N-PSDB; ABX15947, ABX15948.

XX New zidnt5 polypeptides, useful for diagnosing, preventing or treating

PT tumor formation, Crohn's disease, inflammatory bowel disease, food

PT poisoning, melanoma, degenerative disease, bacterial infection, or for

PT wound healing.

XX Claim 2; Page 32-34; 37pp; English.

XX The invention relates to an isolated polypeptide, zidnt5, (a disintegrin

CC and metalloprotease (ADAM), the gene for which is located on human

CC chromosome 9q34), comprising the metalloprotease domain, thrombospondin-

CC 1 domains (TSP-1) and mature protein. Also included are an isolated

CC polynucleotide encoding a fusion protein comprising a first and a second

CC polypeptide segments (where the first polypeptide segment comprises the

CC polypeptide with the sequence of S2, and the second polypeptide comprises

CC one or more (TSP1)-like domains, and where the first polynucleotide

CC segment is positioned amino-terminally to the second polynucleotide

CC segment), an expression vector (comprising the following operably linked

CC elements; (a) a transcription promoter; (b) a DNA segment encoding a

CC polypeptide comprising the amino acid sequences listed above; and (c) a

CC transcription terminator), a cultured cell into which the expression

CC vector of has been introduced, modulating extracellular matrix

CC interactions by combining the polypeptides with cells, producing an

CC antibody to the polypeptides and an isolated immunogenic polypeptide

CC comprising at least 30 contiguous amino acids of mature zidnt5. The

CC polypeptides and polynucleotides are useful for modulating cellular

CC interactions. The polypeptides and antibodies to zidnt5 are useful for in

CC vivo diagnostic or therapeutic applications, such as in identifying or

CC treating tissues or organs that express a corresponding anti-

CC complementary molecule. The polypeptides are useful for diagnosing,

CC preventing or treating tumour formation, Crohn's disease, inflammatory

CC bowel disease, food poisoning, melanoma, degenerative disease, bacterial

CC infections, extracellular matrix repair and remodeling, or for wound

CC healing. The polypeptides are also useful for the study of host defence,

CC extracellular matrix repair, proteolysis, apoptosis, angiogenesis and as

CC a detergent. The antibodies may be used for tagging cells that express

CC zidnt5, for isolating zidnt5 by affinity purification, or for diagnostic

CC assays for determining circulating levels of zidnt5 polypeptides. The

CC present sequence represents mature zidnt5

XX

SQ Sequence 1120 AA;

Query Match 92.1%; Score 492; DB 6; Length 1120;

Best Local Similarity 100.0%; Pred. No. 2.9e-42;

Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 KAPSHSAPLLGLALLRMHQHPRARCPPLCVAGILACGFLGCGWGPSHFQOSCLQALEPQ 67

Db 15 KAPSHSAPLLGLALLRMHQHPRARCPPLCVAGILACGFLGCGWGPSHFQOSCLQALEPQ 74

Qy 68 AVSSYLSPGAPLKGRRPPSPGFQQRQRR 97

Db 75 AVSSYLSPGAPLKGRRPPSPGFQQRQRR 104

RESULT 4

AAE24450

ID AAE24450 standard; protein; 1427 AA.

XX AAE24450;

XX 04-OCT-2002 (first entry)

DT Human Von Willebrand factor-cleaving protease (vWF-cp).

DE Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;

XX transgenic animal; immunisation; thromboembolic disease; preeclampsia;

KW thrombotic thrombocytopenic purpura; TTP; Henoch-Schonlein purpura;

KW thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;

KW transgenic; anticoagulant; chromosome 9.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..51

FT /label= Signal_peptide

FT Protein 52..1427

FT /note= "Mature human vWF-cp protein"

FT Cleavage-site 69..75

FT /note= "Furin cleavage site"

FT Region 224..228

FT /note= "Catalytical side"

FT Region 249

FT /note= "Met turn"

FT Domain 301..377

FT /note= "Disintegrin like motif"

FT Domain 387..439

FT /note= "Thrombospondin type I motif"

FT Region 441..553

FT /note= "Cysteine rich region"

FT Region 554..687

FT /note= "Spacer"

FT Domain 688..743

FT /note= "Thrombospondin type I motif"

FT Domain 744..805

FT /note= "Thrombospondin type I motif"

FT Domain 897..952

FT /note= "Thrombospondin type I motif"
 FT 953..1013
 FT /note= "Thrombospondin type I motif"
 FT 1016..1073
 FT /note= "Thrombospondin type I motif"
 FT 1075..1131
 FT /note= "Thrombospondin type I motif"
 XX
 PN WO200242441-A2.
 XX
 PD 30-MAY-2002.
 XX
 XX 20-NOV-2001; 2001WO-EP013391.
 PF
 XX 22-NOV-2000; 2000US-00721254.
 PR
 PR 12-APR-2001; 2001US-00833328.
 XX
 XX (BAXT) BAXTER AG.
 PA
 XX Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
 PI Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
 PI Zimmermann K, Voelkel D;
 XX
 XX WPI; 2002-479950/51.
 DR
 DR N-PSDB; AAD39332.
 XX
 XX Novel isolated or substantially purified Von Willebrand factor-cleaving
 PT protease, useful for producing preparation for therapy of thrombosis and
 PT thromboembolic disease such as thrombotic thrombocytopenic purpura.
 PT
 XX Claim 1; Fig 5; 93pp; English.
 PS
 XX The invention relates to an isolated or substantially pure Von Willebrand
 CC factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
 CC purifying vWF which involves providing vWF-cp as a ligand, contacting a
 CC solution comprising vWF with the polypeptide ligand under conditions
 CC where vWF is bound to the ligand and recovering from the ligand purified
 CC vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
 CC which involves immunising an animal with vWF-cp and isolating the anti-
 CC vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
 CC producing a preparation of prophylaxis and therapy of thrombosis and
 CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
 CC Henoch-Schönlein purpura, pre-eclampsia, neonatal thrombocytopenia or
 CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing
 CC plasmas or recombinantly produced vWF. The invention is useful for
 CC construction expression systems and generating transgenic animals which
 CC express the polypeptide in vivo. The present sequence is human vWF-cp
 CC protein. vWF-cp gene is located on chromosome 9
 XX
 XX Sequence 1427 AA;
 SQ
 Query Match 77.5%; Score 414; DB 5; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 4.7e-34;
 Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 24 MQQRHPRARCPPLCVAGILACGFLGCGWGPSHFQSCLOALEPQAVSSYLSPGAPLKGRP 83
 Db 1 MQQRHPRARCPPLCVAGILACGFLGCGWGPSHFQSCLOALEPQAVSSYLSPGAPLKGRP 60
 QY 84 PSPGFQRQRQR 97
 Db 61 PSPGFQRQRQR 74
 RESULT 5
 AAD94038
 ID AAD94038 standard; protein; 1427 AA.
 XX
 AC AAD94038;
 XX
 XX 29-JAN-2004 (first entry)
 DT
 XX Human aggrecanase protein amino acid sequence.
 DE

XX /note= "Thrombospondin type I motif"
 XX 953..1013
 XX /note= "Thrombospondin type I motif"
 XX 1016..1073
 XX /note= "Thrombospondin type I motif"
 XX 1075..1131
 XX /note= "Thrombospondin type I motif"
 OS
 XX Homo sapiens.
 XX
 PN US2003105313-A1.
 XX
 PD 05-JUN-2003.
 XX
 XX 25-JAN-2002; 2002US-00057487.
 PF
 XX 16-OCT-2001; 2001US-00978979.
 PR
 XX (AMHP) AMERICAN HOME PROD CORP.
 PA
 XX Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
 PI WPI; 2003-801251/75.
 XX
 DR N-PSDB; ADD94037.
 DR
 XX New isolated DNA molecule encoding an aggrecanase polypeptide for
 PT producing a purified human aggrecanase protein which can be used to
 PT develop inhibitors of aggrecanase.
 PT
 XX Claim 14; SEQ ID NO 8; 24pp; English.
 PS
 XX This invention relates to a novel human aggrecanase protein and the DNA
 CC sequence which encodes it. Aggrecan is a major extracellular component of
 CC articular cartilage. It is a proteoglycan responsible for providing
 CC cartilage with its mechanical properties of compressibility and
 CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
 CC cleavage of aggrecan thereby having a role in cartilage degradation
 CC associated with osteoarthritis and inflammatory joint disease. Compounds
 CC which inhibit the activity of the protein of the invention may have
 CC antiarthritic, osteopathic or antiinflammatory activity. The invention
 CC may be used to produce a purified human aggrecanase protein. The protein
 CC (or fragment) may be used to develop inhibitors of aggrecanase, using
 CC three dimensional structural analysis or computer aided drug design. A
 CC peptide which binds to aggrecanase is used to inhibit the proteolytic
 CC degradation of aggrecan. The invention may be useful for the development
 CC of therapeutics for the treatment of aggrecanase-associated disorders,
 CC such as, osteoarthritis and other inflammatory conditions. The present
 CC sequence is that of the human aggrecanase protein (full length sequence)
 CC of the invention.
 XX
 XX Sequence 1427 AA;
 SQ
 Query Match 76.0%; Score 406; DB 7; Length 1427;
 Best Local Similarity 98.6%; Pred. No. 3.2e-33;
 Matches 73; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 24 MQQRHPRARCPPLCVAGILACGFLGCGWGPSHFQSCLOALEPQAVSSYLSPGAPLKGRP 83
 Db 1 MQQRHPRARCPPLCVAGILACGFLGCGWGPSHFQSCLOALEPQAVSSYLSPGAPLKGRP 60
 QY 84 PSPGFQRQRQR 97
 Db 61 PSPGFQRQRQR 74
 RESULT 6
 AAU97641
 ID AAU97641 standard; protein; 242 AA.
 XX
 AC AAU97641;
 XX
 XX 27-AUG-2002 (first entry)
 DT
 XX Human aggrecanase protein.
 DE

XX Aggrecanase; osteopathic; antiarthritic; antiinflammatory; human;
 KW cartilage; osteoarthritis; inflammatory disease; enzyme.
 XX Homo sapiens.
 OS WO200233093-A2.
 PN 25-APR-2002.
 XX 17-OCT-2001; 2001WO-US032458.
 XX 18-OCT-2000; 2000US-0241469P.
 XX (GEMY) GENETICS INST INC.
 PA Racie LA, Twine NC, Agostino MJ, Wolfman NM, Morris EA;
 PI WPI; 2002-454602/48.
 XX N-PSDB; ABK52579.
 DR Novel purified aggrecanase polypeptide useful for developing inhibitors
 XX and antibodies to the aggrecanase polypeptide, which are useful for
 PT treating aggrecanase-associated condition such as osteoarthritis.
 PT Claim 11; Page 33-34; 41pp; English.
 XX This invention relates to the cDNA and protein sequences of a novel human
 CC aggrecanase polypeptide. The protein of the invention may be used to
 CC inhibit the proteolytic activity of aggrecanase, or to inhibit the
 CC aggrecanase-mediated cleavage of aggrecan in cartilage. The protein of
 CC the invention is useful for developing inhibitors of aggrecanase protein.
 CC The cDNA sequence encoding the aggrecanase protein of the invention is
 CC useful for designing probes for obtaining DNA sequences encoding other
 CC aggrecanase molecules. The cDNA sequence is also useful for detecting
 CC mRNA encoding aggrecanase in a given cell population, and thus for
 CC detecting or diagnosing genetic disorders involving the aggrecanase, or
 CC disorders involving cellular, organ or tissue disorders in which
 CC aggrecanase is irregularly transcribed or expressed. The DNA sequences
 CC may also be useful for preparing vectors for gene therapy applications.
 CC An inhibitor of the protein is useful in treating conditions
 CC characterized by degradation of articular cartilage, by blocking the
 CC enzyme's proteolytic activity. An aggrecanase protein inhibitor and a
 CC method for inhibition of its activity are useful for treating various
 CC aggrecanase-associated conditions including osteoarthritis and other
 CC inflammatory diseases. The present sequence represents the human
 CC aggrecanase protein of the invention
 XX Sequence 242 AA;
 SQ
 Query Match 38.2%; Score 204; DB 5; Length 242;
 Best Local Similarity 100.0%; Pred. No. 4.5e-13;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 SCLQALEPQAVSSYLSFGAPLKGRRPPSPGFQQRQRR 97
 Db 3 SCLQALEPQAVSSYLSFGAPLKGRRPPSPGFQQRQRR 41
 RESULT 7
 ADD94031
 ID ADD94031 standard; protein; 242 AA.
 XX ADD94031;
 AC 29-JAN-2004 (first entry)
 XX Human aggrecanase protein partial amino acid sequence.
 DE aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
 XX cartilage degradation; osteoarthritis; inflammatory joint disease;
 KW antiarthritic; osteopathic; antiinflammatory;
 KW aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
 XX

KW human; enzyme.
 XX Homo sapiens.
 XX US2003105313-A1.
 PN 05-JUN-2003.
 XX 25-JAN-2002; 2002US-00057487.
 XX 16-OCT-2001; 2001US-00978979.
 XX (AMHP) AMERICAN HOME PROD CORP.
 PA Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
 PI WPI; 2003-801251/75.
 XX N-PSDB; ADD94032, ADD94033.
 DR New isolated DNA molecule encoding an aggrecanase polypeptide for
 XX producing a purified human aggrecanase protein which can be used to
 PT develop inhibitors of aggrecanase.
 PT Claim 13; SEQ ID NO 1; 24pp; English.
 XX This invention relates to a novel human aggrecanase protein and the DNA
 CC sequence which encodes it. Aggrecan is a major extracellular component of
 CC articular cartilage. It is a proteoglycan responsible for providing
 CC cartilage with its mechanical properties of compressibility and
 CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
 CC cleavage of aggrecan thereby having a role in cartilage degradation
 CC associated with osteoarthritis and inflammatory joint disease. Compounds
 CC which inhibit the activity of the protein of the invention may have
 CC antiarthritic, osteopathic or antiinflammatory activity. The invention
 CC may be used to produce a purified human aggrecanase protein. The protein
 CC (or fragment) may be used to develop inhibitors of aggrecanase, using
 CC three dimensional structural analysis or computer aided drug design. A
 CC peptide which binds to aggrecanase is used to inhibit the proteolytic
 CC degradation of aggrecan. The invention may be useful for the development
 CC of therapeutics for the treatment of aggrecanase-associated disorders,
 CC such as osteoarthritis and other inflammatory conditions. The present
 CC sequence is that of the human aggrecanase protein (partial) of the
 CC invention.
 XX Sequence 242 AA;
 SQ
 Query Match 38.2%; Score 204; DB 7; Length 242;
 Best Local Similarity 100.0%; Pred. No. 4.5e-13;
 Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 59 SCLQALEPQAVSSYLSFGAPLKGRRPPSPGFQQRQRR 97
 Db 3 SCLQALEPQAVSSYLSFGAPLKGRRPPSPGFQQRQRR 41
 RESULT 8
 AAU79217
 ID AAU79217 standard; protein; 933 AA.
 XX AAU79217;
 AC 15-JUL-2002 (first entry)
 XX Human ADAM-TS-like protein.
 DT Human; ADAM-TS-like protein; cardiovascular disorder; angina;
 XX vascular system; congestive heart failure; myocardial infarction;
 KW ischaemic heart disease; arrhythmia; hypertensive vascular disease;
 KW secondary arterial hypertension; peripheral vascular disease; embolism;
 KW chronic peripheral arterial occlusive disease; acute arterial thrombosis;
 KW inflammatory vascular disorder; chronic obstructive pulmonary disease;
 KW liver disorder.
 XX

[illegible]

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation.
 XX Homo sapiens.
 OS WO200164835-A2.
 PN 07-SEP-2001.
 PD 26-FEB-2001; 2001WO-US004927.
 PF 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 PS (HYSE-) HYSEQ INC.
 PP Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-514838/56.
 XX N-PSDB; AA186390.
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 XX Claim 20; SEQ ID NO 20351; 1399pp + Sequence Listing; English.
 XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA00010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 115 AA;
 SQ
 Query Match 17.4%; Score 93; DB 4; Length 115;
 Best Local Similarity 27.1%; Pred. No. 0.069;
 Matches 29; Conservative 17; Mismatches 39; Indels 22; Gaps 4;
 QY 4 PTPKAPSHSAPLLGLALLRMHQRHPRAR---CPPLCVAGILACQFL---GCWGFSHFQ 57
 Db 8 PRFRKESGPPPPKVFLETRQPGQTKPPFPPI-----FKLTRPGRRPXPFP 58
 QY 58 QSCLOALEPOAVSSYLSFGA-----PLKRRPSPGPRQRQR 97
 Db 59 PORVRAGKPPXPOSFRVPGAIFPPPPPPKLGHPRRPGQKKKKERK 105
 RESULT 11
 ABP98888
 ID ABP98888 standard; protein; 938 AA.
 AC ABP98888;
 DT 24-JUL-2003 (first entry)
 XX Human molecule for disease detection and treatment MDDT-16.
 DE Cytostatic; antiarteriosclerotic; anti-HIV; antiallergic; nephrotropic;
 KW antithyroid; cerebroprotective; antiparkinsonian; anticonvulsant; MDDT;
 KW neurotropic; neuroprotective; antidiabetic; gene therapy; atherosclerosis;
 KW molecule for disease detection and treatment; cancer; AIDS; allergy;
 KW diabetes; glomerulonephritis; autoimmune thyroiditis; Cushing's syndrome;
 KW stroke; Parkinson's disease; epilepsy.

XX Homo sapiens.
 OS WO2003031595-A2.
 PN 17-APR-2003.
 PD 10-OCT-2002; 2002WO-US032852.
 PF 12-OCT-2001; 2001US-0328944P.
 PR 26-OCT-2001; 2001US-0345384P.
 PR 02-NOV-2001; 2001US-0343880P.
 PR 09-NOV-2001; 2001US-0345143P.
 PR 16-NOV-2001; 2001US-0332430P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Tang YT, Forsythe IJ, Emerling BM, Hafalia AJA, Yue H, Xu Y;
 PI Gietzen KJ, Chawla NK, Baughn MR, Marquis JP, Becha SD, Kable AE;
 PI Lal PG, Richardson TW, Lee SY, Lee EA, Tran B, Warren BA, Lu DAM;
 PI Gururajan R, Sprague WW, Blake JJ, Thangavelu K, Swarnakar A;
 PI Gorvad AE, Griffin JA, Lindquist BA, Elliott VS, Ison CH;
 PI Ramkumar J;
 XX WPI; 2003-421277/39.
 DR N-PSDB; ACC44403.
 XX Isolated peptide molecules for disease detection and treatment, useful
 PT for diagnosing, treating or preventing disorders, e.g. cancer, AIDS,
 PT atherosclerosis, diabetes or stroke.
 XX Claim 1; Page 163-165; 234pp; English.
 XX The invention relates to the isolation of a number of "molecules for
 CC disease detection and treatment" (MDDT) and genes encoding them. The
 CC invention also includes molecule which are at least 90% identical to the
 CC protein and nucleotide sequences. This sequence represents a protein of
 CC the invention. Disorders associated with aberrant expression of MDDT, are
 CC cell proliferative disorders (e.g. cancer or atherosclerosis),
 CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, diabetes,
 CC glomerulonephritis or autoimmune thyroiditis), developmental disorders
 CC (e.g. Cushing's syndrome) or neurological disorders (e.g. stroke,
 CC Parkinson's disease or epilepsy)
 XX Sequence 938 AA;
 SQ
 Query Match 16.2%; Score 86.5; DB 6; Length 938;
 Best Local Similarity 24.8%; Pred. No. 2.9;
 Matches 28; Conservative 10; Mismatches 38; Indels 37; Gaps 5;
 QY 2 GRPTRKAPSHSAPLLGLALLRMHQRHPRAR-----CPPLC-----V 38
 Db 258 GPCCQDTGPTHYPP-----PHPPPPHPPQALPCPPACRHPKQGSYSPALPQL 307
 QY 39 AGILACQFLGCGWSPSHFQSCLOA--LEPOAVSSYLSFGAPLKGRPPSPGQ 89
 Db 308 GGHKGITGYAGGLGSPVLRQQAQAPYIPPLGLDAYPSPALPA--PSPGLK 358
 RESULT 12
 ABG31634
 ID ABG31634 standard; protein; 1047 AA.
 XX ABG31634;
 AC ABG31634;
 XX 29-NOV-2002 (first entry)
 XX Human PPIP-120 protein.
 DE Human; protein histidine phosphatase interacting partner of 120kD;
 KW PPIP-120; autism; anaemia; malignant fibrous histiocytoma; vaccine;
 KW cyostatic; nontropic; antianaemic; PPIP-120-Agonist; gene therapy;
 KW PPIP-120-Antagonist.

```

XX OS Homo sapiens.
XX PN WO200266507-A2.
XX XX
XX PD 29-AUG-2002.
XX XX
XX PF 25-JAN-2002; 2002WO-EP000753.
XX PR 16-FEB-2001; 2001EP-00103779.
XX PA (MERE ) MERCK PATENT GMBH.
XX PI Hock B, Duecker K, Kellner R;
XX XX
XX DR WPI; 2002-674915/72.
XX DR N-PSDB; ABS53396.
XX XX
XX PT New Protein Histidine Phosphatase Interacting Partner of 120kD ligand
XX PT polypeptides, useful for diagnosing or treating diseases e.g. autism,
XX PT anemia or malignant fibrous histiocytomas.
XX PS Claim 2; Page 36-38; 41pp; English.
XX CC The present invention relates to a new Protein Histidine Phosphatase
XX CC Interacting Partner of 120kD (PHPIP-120) ligand polypeptide. The
XX CC polynucleotide is useful for chromosome localisation studies or for
XX CC tissue expression studies. The polypeptide is useful for treating
XX CC diseases e.g. autism, anaemia or malignant fibrous histiocytomas. The
XX CC polynucleotide and polypeptide are also useful in diagnostic assays or as
XX CC vaccines against the mentioned diseases. The present amino acid sequence
XX CC represents the human PHPIP-120 protein of the invention
XX XX
XX SQ Sequence 1047 AA;
XX
XX Query Match 16.2%; Score 86.5; DB 5; Length 1047;
XX Best Local Similarity 24.8%; Pred. No. 3.3;
XX Matches 28; Conservative 10; Mismatches 38; Indels 37; Gaps 5;
XX
QY 2 GRPTRKAPSHSAPLLGLALLRMHQRHPRAR-----CPPLC-----V 38
DB 258 GPCQGTGTHYPP-----PHPPPHPPQALPCPPACRHPKQGSYSPALPLQPL 307
QY 39 AGILACGLLGCWGPSHFQSCLOA--LEPQAVSSYLSFGAPLKGRRPSPGFQ 89
DB 308 GGHKGTGYAGGLGSPYLQQAQAQAPYIPPLGLDAYPYPSAPLPA--RSPGLK 358
XX
RESULT 13
ID ADC32956
XX ADC32956 standard; protein; 539 AA.
XX AC ADC32956;
XX XX
XX DT 18-DEC-2003 (first entry)
XX XX
XX DE Human novel contig-encoded polypeptide sequence, SRQ ID NO:3038.
XX XX
XX KW Human; diagnostic; drug screening; forensics; gene mapping;
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX KW ulcers; osteoporosis; autoimmune disease; cancer;
XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
XX KW neuroprotective; anti-anaemic; anticoagulant; thrombolytic; vulnerary;
XX KW anti-ulcer; osteopathic; immunosuppressive; anti-inflammatory; cytostatic;
XX KW gene therapy; chromosome 16.
XX OS
XX OS Homo sapiens.
XX PN WO2003029271-A2.
XX XX
XX PD 10-APR-2003.
XX XX

```

```

PF 24-SEP-2002; 2002WO-US030474.
XX XX
XX PR 24-SEP-2001; 2001US-0324631P.
XX XX
XX PA (HYSE-) HYSEQ INC.
XX XX
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX PI Haley-Vicente D, Drmanac RT;
XX XX
XX DR WPI; 2003-371981/35.
XX DR N-PSDB; ADC32189.
XX XX
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
XX PT treating conditions such as neurodegenerative diseases, anemias, platelet
XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX PT cancer.
XX PS
XX XX
XX Example 2; SEQ ID NO 3038; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX invention also relates to nucleic acid sequences over 99% identical with
XX the novel human cDNAs. The invention additionally encompasses expression
XX vectors and host cells comprising a nucleic acid of the invention; the
XX recombinant production of a polypeptide of the invention; an antibody
XX against a polypeptide of the invention; a method of detecting
XX polynucleotides or polypeptides of the invention; and methods of
XX identifying a compound which binds to a polypeptide of the invention. The
XX invention further discloses methods of preventing, treating or
XX ameliorating a medical condition; kits comprising polynucleotide probes
XX and/or monoclonal antibodies for carrying out the methods of the
XX invention; methods for the identification of compounds that modulate the
XX expression or activity of the polynucleotide and/or polypeptide; and 767
XX contig sequences corresponding to the cDNA sequences of the invention
XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
XX -ADC33394). The nucleic acids and polypeptides of the invention are
XX useful in diagnostics, drug screening, forensics, gene mapping, in the
XX identification of mutations responsible for genetic disorders or other
XX traits, for assessing biodiversity, and in producing many other types of
XX data and products dependent on DNA and amino acid sequences. They are
XX also used for treating diseases such as Parkinson's disease, Alzheimer's
XX disease and other neurodegenerative diseases, anaemia, platelet
XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer. The nucleic acids may also be used as hybridisation probes or
XX primers, and in the recombinant production of a protein. The polypeptides
XX are also useful in generating antibodies, as molecular weight markers,
XX and as food supplements. The present sequence represents a human contig-
XX encoded polypeptide sequence used in an example of the invention. Note:
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX XX
XX SQ Sequence 539 AA;
XX
XX Query Match 15.2%; Score 81; DB 7; Length 539;
XX Best Local Similarity 32.5%; Pred. No. 6.1;
XX Matches 40; Conservative 7; Mismatches 34; Indels 42; Gaps 9;
XX
QY 1 PGPTPRKAPSHSAPLLGLALLRMHQRHPRARCP--LC----- 37
DB 350 PG-PSRP--PSX---LQGLASVPAGHPPRALGSGGPGCPDXDGATVUSDPEGEAGAE 403
QY 38 ---VAGILACG--FLLGCWGPSHFQSCLOALE--PQAVSSYLS--PQAPLKGRRPSPGF 88
DB 404 REEVAGDLACGNLTLDGWSFX----SCVPSEEDSEQVAMCSPGCGPEQAPGAAGR 459
QY 89 QRQ 91
DB 460 QRQ 462
XX
RESULT 14

```

```
ABP76680
ID ABP76680 standard; protein; 19938 AA.
XX AC
XX ABP76680;
XX
DT 26-FEB-2003 (first entry)
XX
XX Streptomyces viridochromogenes Avi gene cluster polypeptide frame 4.
XX
XX Avilamycin; antibacterial; virucide; protozoacide; fungicide; infection;
KW medicine; Staphylococcus aureus; biosynthetic gene cluster; enzyme.
XX
XX Streptomyces viridochromogenes.
OS
XX WO200268436-A1.
PN
XX 06-SEP-2002.
PD
XX
XX 24-AUG-2001; 2001WO-EP009815.
PF
XX
XX 25-FEB-2001; 2001DE-01009166.
PR
XX (COMB-) COMBINATURE BIOPHARM AG.
PA
XX Weitnauer G, Muehlenweg A, Trefzer A, Bechthold A;
PI
XX WPI; 2003-018650/01.
DR
XX N-PSDB; ABZ37516.
DR
XX
XX New avilamycin derivatives, useful for treatment of infections, and
PT nucleic acid encoding avilamycin synthesis enzymes.
PT
XX
XX Example 1; Page 68-301; 319pp; German.
PS
XX
XX The invention relates to avilamycin derivatives (I) with antibacterial,
CC virucide, protozoacide and fungicide activity. (I) are useful for
CC treatment of infections (bacterial, viral, protozoal or fungal), in human
CC or veterinary medicine, particularly where caused by Staphylococcus
CC aureus. (II) are more hydrophilic than known avilamycins. The present
CC sequence is that of an avilamycin synthesis enzyme from the Streptomyces
CC viridochromogenes Avilamycin A biosynthetic gene cluster (ABZ37515-
CC ABZ37516)
CC
XX
XX Sequence 19938 AA;
SQ
Query Match 15.1%; Score 80.5; DB 6; Length 19938;
Best Local Similarity 25.8%; Pred. No. 3e+02;
Matches 32; Conservative 5; Mismatches 28; Indels 59; Gaps 4;
QY 1 PGRPTRKAPSHSAPLL---GLALLRMHQHPRARCPLCVAGILACGFLGCGWGFHFQ 57
Db 4633 PPRPARPPPSRTARRAARGARAAPRAASRPFRPPRC----- 4672
QY 58 QSCQLALEPQAVSYLSGAPL-----KGPP-----SPGFQRQRQ 93
Db 4673 -----GRRRSFGAPRPGRRGHRCPGSSPRCATTGRRGPPGTRRSPGRRRR 4720
QY 94 RGR 97
Db 4721 PGR 4724
RESULT 15
AA001892
ID AAC001892 standard; protein; 202 AA.
XX
XX AAC001892;
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 15784.
DE
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
```

```
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
XX N-PSDB; AAI81823.
DR
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 20; SEQ ID NO 15784; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO3910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 202 AA;
SQ
Query Match 15.0%; Score 80; DB 4; Length 202;
Best Local Similarity 25.9%; Pred. No. 2.8;
Matches 37; Conservative 8; Mismatches 36; Indels 62; Gaps 7;
QY 4 PTPKAPSHSAPLLGL-----ALLRMHQH--HPRARCP-----ELCVAG 40
Db 16 PKLPQAQLSSPGLNLVFTAGSNDAGLLRLSPRANRVALPRVHTGSSSTLSPSCPLG 75
QY 41 ILACGF-----LLCGWGFHFQSCICQ--A 63
Db 76 -LACGFTSQPACIXSPFLLAPARPSGCAWPGLPVCSLHLRGWGGAD-PTGCLDAWA 133
QY 64 LEQAVSSYLSGAPLKGRRPSP 86
Db 134 MALASLRPCLCPALPLSRPPSP 156
Search completed: March 13, 2004, 07:39:05
Job time : 17.1964 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 7.67917 Seconds
(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534
Sequence: 1 PGRPTRPKAPSHSAPLLGLA.....PLKGRPPSPGFQFQRRQRR 97

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	534	100.0	1416	9	US-09-836-712-2
2	492	92.1	1120	9	US-09-781-080B-11
3	414	77.5	842	14	US-10-222-334-4
4	414	77.5	1427	14	US-10-222-334-2
5	406	76.0	1427	14	US-10-057-487-8
6	204	38.2	242	9	US-09-978-979-1
7	204	38.2	242	14	US-10-057-487-1
8	84	15.7	760	15	US-10-369-493-12509
9	80.5	15.1	19662	15	US-10-084-846A-6
10	79	14.8	117	14	US-10-097-111-340
11	79	14.8	412	15	US-10-264-049-3139
12	78.5	14.7	145	15	US-10-264-049-3571
13	77	14.4	461	14	US-10-106-698-4705
14	76	14.2	216	9	US-09-893-737-290
15	74.5	14.0	175	9	US-09-764-877-1158

16	74.5	14.0	175	9	US-09-860-670-104	Sequence 104, Appl
17	74.5	14.0	175	15	US-10-227-646-104	Sequence 104, Appl
18	74.5	14.0	175	15	US-10-242-515-1158	Sequence 1158, Ap
19	74.5	14.0	1973	15	US-10-369-493-2380	Sequence 2380, Ap
20	74	13.9	4123	14	US-10-213-509-5	Sequence 5, Appli
21	74	13.9	4219	15	US-10-085-198-2	Sequence 2, Appli
22	73.5	13.8	274	15	US-10-108-260A-3891	Sequence 3891, Ap
23	73	13.7	753	9	US-09-908-180-2	Sequence 2, Appli
24	72.5	13.6	439	14	US-10-156-761-12284	Sequence 12284, A
25	72.5	13.6	737	14	US-10-017-161-1350	Sequence 1350, Ap
26	72.5	13.6	737	15	US-10-292-798-1110	Sequence 1110, Ap
27	72.5	13.6	1021	14	US-10-017-161-740	Sequence 740, App
28	72	13.5	325	15	US-10-295-027-1374	Sequence 1374, Ap
29	72	13.5	325	15	US-10-367-978-23	Sequence 23, Appl
30	72	13.5	326	9	US-09-925-301-1349	Sequence 1349, Ap
31	71.5	13.4	203	15	US-10-108-260A-2454	Sequence 2454, Ap
32	71.5	13.4	19695	15	US-10-084-846A-3	Sequence 3, Appli
33	71	13.3	160	15	US-10-104-047-3106	Sequence 3106, Ap
34	71	13.3	213	15	US-10-094-749-2061	Sequence 2061, Ap
35	71	13.3	618	14	US-10-406-209-5	Sequence 5, Appli
36	71	13.3	624	9	US-09-925-297-738	Sequence 738, App
37	71	13.3	624	14	US-10-106-698-4541	Sequence 4541, Ap
38	71	13.3	913	9	US-09-223-490-4	Sequence 4, Appli
39	71	13.3	913	10	US-09-355-815-2	Sequence 2, Appli
40	71	13.3	913	11	US-09-236-939-4	Sequence 4, Appli
41	71	13.3	913	15	US-10-116-275-156	Sequence 156, App
42	71	13.3	913	15	US-10-407-365-4	Sequence 4, Appli
43	71	13.3	919	15	US-10-407-365-6	Sequence 6, Appli
44	70.5	13.2	124	15	US-10-047-021-76	Sequence 76, Appl
45	70.5	13.2	125	14	US-10-411-224-76	Sequence 76, Appl

ALIGNMENTS

RESULT 1

US-09-836-712-2
; Sequence 2, Application US/09836712
; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PCT0851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1416
; TYPE: PRT
; ORGANISM: Human
US-09-836-712-2

Query Match 100.0%; Score 534; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 5e-43;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	PGKPTRPKAPSHSAPLLGLALLRMHQHPHRCPPCLCVAGILACGFLGCGWSPSHFQOSC	60
DB	1	PGKPTRPKAPSHSAPLLGLALLRMHQHPHRCPPCLCVAGILACGFLGCGWSPSHFQOSC	60
QY	61	LQALEPQAVSSYLSPGAPLKGRPPSPGFQFQRRQRR	97
DB	61	LQALEPQAVSSYLSPGAPLKGRPPSPGFQFQRRQRR	97

RESULT 2

US-09-781-080B-11
; Sequence 11, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM

APPLICANT: SHEPPARD, PAUL
APPLICANT: YAMAMOTO, GAYLE
TITLE OF INVENTION: Anti-Angiogenic Intestinal Peptides,
TITLE OF INVENTION: zlns5
FILE REFERENCE: 99-82
CURRENT APPLICATION NUMBER: US/09/781,080B
CURRENT FILING DATE: 2002-01-30
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 1120
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1120)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-781-080B-11

Query Match 92.1%; Score 492; DB 9; Length 1120;
Best Local Similarity 100.0%; Pred. No. 4.4e-39;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 KAPSHSAPLGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQ 67
DB 15 KAPSHSAPLGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQ 74

QY 68 AVSSYLSPGAPLKGRPPSPGFORQORRR 97
DB 75 AVSSYLSPGAPLKGRPPSPGFORQORRR 104

RESULT 3
US-10-222-334-4
Sequence 4, Application US/10222334
Publication No. US20030073116A1
GENERAL INFORMATION:
APPLICANT: Ginsburg, David
APPLICANT: Levy, Gallia
APPLICANT: Tsai, Han-Mou
TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
FILE REFERENCE: UM-07288
CURRENT APPLICATION NUMBER: US/10/222,334
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 60/312,834
PRIOR FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 842
TYPE: PRT
ORGANISM: Homo sapiens
US-10-222-334-4

Query Match 77.5%; Score 414; DB 14; Length 842;
Best Local Similarity 100.0%; Pred. No. 1.1e-31;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 83
DB 1 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 60

QY 84 PSPGFORQORRR 97
DB 61 PSPGFORQORRR 74

RESULT 4
US-10-222-334-2
Sequence 2, Application US/10222334
Publication No. US20030073116A1
GENERAL INFORMATION:
APPLICANT: Ginsburg, David

APPLICANT: Levy, Gallia
APPLICANT: Tsai, Han-Mou
TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
FILE REFERENCE: UM-07288
CURRENT APPLICATION NUMBER: US/10/222,334
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 60/312,834
PRIOR FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 1427
TYPE: PRT
ORGANISM: Homo sapiens
US-10-222-334-2

Query Match 77.5%; Score 414; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.9e-31;
Matches 74; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 83
DB 1 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 60

QY 84 PSPGFORQORRR 97
DB 61 PSPGFORQORRR 74

RESULT 5
US-10-057-487-8
Sequence 8, Application US/10057487
Publication No. US20030105313A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: Aggreacanase Molecules
FILE REFERENCE: 08702-0073
CURRENT APPLICATION NUMBER: US/10/057,487
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 60/241,469
PRIOR FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 1427
TYPE: PRT
ORGANISM: homo sapiens
US-10-057-487-8

Query Match 76.0%; Score 406; DB 14; Length 1427;
Best Local Similarity 98.6%; Pred. No. 1.1e-30;
Matches 73; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 24 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 83
DB 1 MHQRHPRARCPPLCVAGILACGFLGCGWPSHFQOQSCLOALEPQAVSSYLSPGAPLKGRP 60

QY 84 PSPGFORQORRR 97
DB 61 PSPGFORQORRR 74

RESULT 6
US-09-978-979-1
Sequence 1, Application US/09978979
Patent No. US20020151702A1
GENERAL INFORMATION:
APPLICANT: Racie, Lisa, A.
Twine, Natalie, C.
Agostino, Michael, J.
Welfman, Neil
Morris, Elisabeth
TITLE OF INVENTION: Aggreacanase Molecules

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: American Home Products Corporation
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICANT: US/09/978,979
FILING DATE: 16-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: GI 5435p
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-978-979-1

Query Match 38.2%; Score 204; DB 9; Length 242;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 SCLQALEPQAVSYLSPGAPLKGPPSPGFORQORRR 97
DB 3 SCLQALEPQAVSYLSPGAPLKGPPSPGFORQORRR 41

RESULT 7
US-10-057-487-1
Sequence 1, Application US/10057487
Publication No. US20030105313A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: Aggrecanase Molecules
FILE REFERENCE: 08702-0073
CURRENT APPLICATION NUMBER: US/10/057,487
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 60/241,469
PRIOR FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 242
TYPE: PRT
ORGANISM: Homo sapiens
US-10-057-487-1

Query Match 38.2%; Score 204; DB 14; Length 242;
Best Local Similarity 100.0%; Pred. No. 5.6e-12;
Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 59 SCLQALEPQAVSYLSPGAPLKGPPSPGFORQORRR 97
DB 3 SCLQALEPQAVSYLSPGAPLKGPPSPGFORQORRR 41

RESULT 8
US-10-369-493-12509
Sequence 12509, Application US/10369493
Publication No. US20030233675A1
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Goldman, Barry S.
APPLICANT: Chen, Xianfeng
TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
FILE REFERENCE: 38-10(52052)B
CURRENT APPLICATION NUMBER: US/10/369,493
CURRENT FILING DATE: 2003-02-28
PRIOR APPLICATION NUMBER: US 60/360,039
PRIOR FILING DATE: 2002-02-21
NUMBER OF SEQ ID NOS: 47374
SEQ ID NO 12509
LENGTH: 760
TYPE: PRT
ORGANISM: Aspergillus nidulans
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(760)
OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12509

Query Match 15.7%; Score 84; DB 15; Length 760;
Best Local Similarity 26.6%; Pred. No. 6.9;
Matches 33; Conservative 11; Mismatches 40; Indels 40; Gaps 5;
QY 1 PGRPTRPKAP-----SHSAPLL-----GLALLRMHQHPRARCPPLCVAG 40
DB 284 PQSPTSRQQLPRTSSSPAAAAAATGMPGKXQGLALLRRYHLKXNCRPTCSAP 343
QY 41 ILACGELLSCWGSHPQSCLOALEPQAVSYL--SGAPLKGPP-----PSPGFORQ 91
DB 344 -----GPCSLRPPCCIRILPARRYAARLPASPGNAEGNAPAQYGMHAPPPPPQQQ 392
QY 92 RORQ 95
DB 393 QOMQ 396

RESULT 9
US-10-084-846A-6
Sequence 6, Application US/10084846A
Publication No. US20040006026A1
GENERAL INFORMATION:
APPLICANT: WEITNAUER, GABRIELE
APPLICANT: MUHLENWEG, AGNES
APPLICANT: TREPZER, AXEL
APPLICANT: BECHTHOLD, ANDREAS
TITLE OF INVENTION: AVILAMYCIN DERIVATIVES
FILE REFERENCE: 1974-005
CURRENT APPLICATION NUMBER: US/10/084,846A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: PCT/EP01/09815
PRIOR FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: DE 101 09 166.4
PRIOR FILING DATE: 2001-02-25
NUMBER OF SEQ ID NOS: 120
SOFTWARE: PatentIn Ver. 3.2
SEQ ID NO 6
LENGTH: 19662
TYPE: PRT
ORGANISM: Streptomyces viridochromogenes
FEATURE:
OTHER INFORMATION: Protein 1: amino acid sequence encoded by coding strand 2.
OTHER INFORMATION: Start codon: gga, Start position: nucleotide 1.

```
US-10-084-846A-6
Query Match      15.1%; Score 80.5; DB 15; Length 19662;
Best Local Similarity 25.8%; Pred. No. 4.4e+02;
Matches 32; Conservative 5; Mismatches 28; Indels 59; Gaps 4;

QY 1 PGRTPRKAPSHSAPILL---GLALLRMHQRHPRARCPPLCVAGILACGFLLCGNGPSHFQ 57
DB 4556 PPRPARPPRPTARRAARGARAPRAASRPRRPPRC-----KGRPP-----SPGFQRQRQ 93
QY 58 QSCLOALEPQAVSSYLSPGAPL-----KGRPP-----SPGFQRQRQ 93
DB 4596 -----GRRSPGAPRPPRRSHRCFGSSPRCAITGRRPFGTRRSPGRRGR 4643
QY 94 QRER 97
DB 4644 QGRR 4647

RESULT 10
US-10-097-111-340
; Sequence 340, Application US/10097111
; Publication No. US20030138771A1
; GENERAL INFORMATION:
; APPLICANT: PELLETIER, JERRY
; APPLICANT: GEOS, PHILLIPPE
; APPLICANT: DUBOW, MICHAEL
; TITLE OF INVENTION: DNA SEQUENCES FROM S. PNEUMONIAE BACTERIOPHAGE EPI THAT
; FILE REFERENCE: 073406-0603
; CURRENT APPLICATION NUMBER: US/10/097,111
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 09/676,412
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/157,218
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 552
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 340
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-10-097-111-340

Query Match      14.8%; Score 79; DB 14; Length 117;
Best Local Similarity 32.6%; Pred. No. 3;
Matches 31; Conservative 10; Mismatches 28; Indels 26; Gaps 5;

QY 1 PGRTPRKAPSHSAPILLGLALLRMHQRHPRARCPPLCVAGILACGFLLCGNGPSHFQSC 60
DB 16 PSTPSRPGAPGKPSPLGFS-SRIHVK-----SGTNSLGLLLVLRTPMYPPDSA 64
QY 61 LQALEPQAVSSYL-----SPGAPLKGRPPSP 86
DB 65 LK-LVPMSSAYLITTWDSFTVSP-----ERTPSP 93

RESULT 11
US-10-264-049-3139
; Sequence 3139, Application US/10264049
; Publication No. US2004000579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: US/10/264,049
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; SOFTWARE: Patentin Ver. 3.1
```

```
; SEQ ID NO 3139
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (36)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3139

Query Match      14.8%; Score 79; DB 15; Length 412;
Best Local Similarity 34.3%; Pred. No. 11;
Matches 24; Conservative 7; Mismatches 37; Indels 2; Gaps 2;

QY 25 HQHPRARCPPLCVAGILA-CGFLLCGNGP-SHFQSCLOALEPQAVSSYLSPGAPLKGR 82
DB 30 HRTRTXRCTPAAVCGSSSRCCNSMSSPGCSHTRHSTCPGLGRADIRPRLSGGSLHWA 89
QY 83 PPSPGFQROR 92
DB 90 PGSPGREGSR 99

RESULT 12
US-10-264-049-3571
; Sequence 3571, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: PCT/US01/18569
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: US/10/264,049
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patentin Ver. 3.1
; SEQ ID NO 3571
; LENGTH: 146
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-3571

Query Match      14.7%; Score 78.5; DB 15; Length 146;
Best Local Similarity 27.3%; Pred. No. 4.3;
Matches 27; Conservative 7; Mismatches 36; Indels 29; Gaps 3;

QY 1 PGRTPRKAPSHSAPILLGLALLRMHQRHPRARCPPLCVAGILACGFLLCGNGPSHFQSC 60
DB 10 PGRPPPGRSPS-----PRRAPRPPKGTSSQAMLVWPGTJL---C 49
QY 61 LQALEPQAV-----SSYLSFGAPLKGRPPSPGFQR 90
DB 50 SQRGPPRPVLGVGTAVAGSCSAFSPHAPAGSQPAAPCQR 88

RESULT 13
US-10-106-698-4705
; Sequence 4705, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypepti
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; SOFTWARE: Patentin Ver. 3.1
```

```

; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4705
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (312)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (406)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4705

Query Match      14.4%; Score 77; DB 14; Length 461;
Best Local Similarity 32.7%; Pred. No. 20;
Matches 32; Conservative 7; Mismatches 43; Indels 16; Gaps 6;

Qy 1 PGRPTRPKAPSHSAPLLGLALLR--MHQRH-----PRARCPPLCVAG---ILACGFLGC 50
Db 365 PRPTSLA-SNTHLLGLRDVRDEFQCHCGATPPCCLPACRPSSWTLSGCPMATSC 423

Qy 51 -WGPSHFQSCLOALEPOAVSYLSFGAPLKGRPFSPG 87
Db 424 GWGSTQQDGLFSVE-----SHPWVLPVLTLPKPPGTG 457

RESULT 14
US-09-893-737-290
; Sequence 290, Application US/09893737
; Patent No. US20020110855A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Presnell, Scott R.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41
; CURRENT APPLICATION NUMBER: US/09/893,737
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 290
; LENGTH: 216
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-893-737-290

Query Match      14.2%; Score 76; DB 9; Length 216;
Best Local Similarity 35.4%; Pred. No. 11;
Matches 28; Conservative 4; Mismatches 25; Indels 22; Gaps 5;

Qy 15 PLLGLALLRMQRHPRARCPPLCVAGILACGFLGCGSHFQ-QSCLOALEPCAVSSYL 73
Db 31 FWCGLKAGRLGLRPP-----PPLCS-----GQSHLAGPGCLP--RQQVLSS-- 69

Qy 74 SPGAPLKGRPPSPGFQQR 92
Db 70 SPGVPGEGLLSAPGFQHR 88

RESULT 15
US-09-764-877-1158
; Sequence 1158, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (98)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (129)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (141)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (145)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-877-1158

Query Match      14.0%; Score 74.5; DB 9; Length 175;
Best Local Similarity 27.3%; Pred. No. 12;
Matches 33; Conservative 9; Mismatches 30; Indels 49; Gaps 6;

Qy 1 PGRPTRPKAPSHSAPLLGLALLRMQRHPRARCPPLCV-----GILACGFL----- 47
Db 8 PGRPTRPVHVTKMSLLSKKFCRLRLQ-----IAWLGDWDLERGLLAFAFWAQLRA 56

Qy 48 -----LCCWGPSHFQOS---CLOALEP-----QAVSSVLSFGAPLKGRPPS 85
Db 57 GHRAGGTGCGAWHFGSGWRGSLASVGPVPNNVSQPFXXSASCLASPHPV---PPS 113

Qy 86 P 86
Db 114 P 114

Search completed: March 13, 2004, 08:07:05
Job time : 8.67917 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:34:04 ; Search time 4.09219 Seconds
(without alignments)
1223.727 Million cell updates/sec

Title: US-09-836-712-2_COPY_1_97

Perfect score: 534

Sequence: 1 PGRTPRKAPSHSAPLLGLA.....PLKGRPPSPGQRQRORRR 97

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*

2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*

3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*

4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*

5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*

6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86.5	16.2	556	US-03-252-991A-22670	Sequence 22670, A
2	84	15.7	485	US-03-252-991A-25176	Sequence 25176, A
3	81	15.2	245	US-03-252-991A-27050	Sequence 27050, A
4	81	15.2	314	US-03-252-991A-30702	Sequence 30702, A
5	80.5	15.1	324	US-03-252-991A-24664	Sequence 24664, A
6	79.5	14.9	510	US-03-252-991A-19528	Sequence 19528, A
7	78	14.6	834	US-03-252-991A-17616	Sequence 17616, A
8	77.5	14.5	273	US-03-252-991A-23651	Sequence 23651, A
9	76	14.2	445	US-03-252-991A-28706	Sequence 28706, A
10	75.5	14.1	330	US-03-252-991A-19363	Sequence 19363, A
11	75	14.0	244	US-03-252-991A-31668	Sequence 31668, A
12	74	13.9	271	US-03-252-991A-17078	Sequence 17078, A
13	74	13.9	366	US-03-252-991A-25602	Sequence 25602, A
14	73.5	13.8	213	US-03-252-991A-29365	Sequence 29365, A
15	73	13.7	286	US-03-252-991A-29952	Sequence 29952, A
16	72.5	13.6	156	US-03-252-991A-23897	Sequence 23897, A
17	72.5	13.6	722	US-03-252-991A-26839	Sequence 26839, A
18	72.5	13.6	860	US-03-252-991A-28607	Sequence 28607, A
19	72	13.5	325	US-03-599-360B-74	Sequence 74, Appl
20	72	13.5	428	US-03-252-991A-24452	Sequence 24452, A
21	72	13.5	691	US-03-252-991A-31413	Sequence 31413, A
22	71.5	13.4	155	US-03-252-991A-17807	Sequence 17807, A
23	71.5	13.4	401	US-03-252-991A-19289	Sequence 19289, A
24	71.5	13.4	508	US-03-252-991A-24350	Sequence 24350, A
25	71.5	13.4	1118	US-03-252-991A-32651	Sequence 32651, A
26	71	13.3	149	US-03-252-991A-22548	Sequence 22548, A
27	71	13.3	210	US-09-489-039A-9923	Sequence 9923, Ap

28 71 13.3 309 4 US-09-252-991A-27042 Sequence 27042, A
29 71 13.3 371 4 US-09-328-352-7132 Sequence 7132, Ap
30 71 13.3 492 4 US-09-352-991A-27853 Sequence 27853, A
31 71 13.3 521 3 US-08-738-168B-5 Sequence 5, Appl
32 71 13.3 722 4 US-09-252-991A-24102 Sequence 24102, A
33 71 13.3 913 1 US-08-445-840-4 Sequence 4, Appl
34 71 13.3 913 3 US-08-170-558-4 Sequence 4, Appl
35 71 13.3 913 3 US-08-447-314-4 Sequence 4, Appl
36 71 13.3 913 3 US-08-445-461-4 Sequence 4, Appl
37 71 13.3 913 4 US-09-140-378A-2 Sequence 2, Appl
38 71 13.3 919 1 US-08-326-343A-2 Sequence 2, Appl
39 70.5 13.2 430 4 US-09-252-991A-32661 Sequence 32661, A
40 70.5 13.2 771 1 US-07-923-876-6 Sequence 6, Appl
41 70 13.1 355 4 US-09-352-991A-18117 Sequence 18117, A
42 70 13.1 513 2 US-08-357-533A-10 Sequence 10, Appl
43 70 13.1 513 2 US-08-459-009-10 Sequence 10, Appl
44 70 13.1 513 2 US-08-300-584-2 Sequence 2, Appl
45 70 13.1 513 3 US-08-459-951-10 Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-252-991A-22670 Application US/09252991A
; Sequence 22670, Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22670
; LENGTH: 556
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22670

Query Match 16.2%; Score 86.5; DB 4; Length 556;

Best Local Similarity 28.0%; Pred. No. 0.21;

Matches 33; Conservative 29; Mismatches 29; Indels 49; Gaps 6;

QY 2 GRTRPKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGF-----LLCGWGP 53

DB 88 GHAPRPLPAPLRP-----VHQH-----APGPRTRGRDQRLARGAQP 125

QY 54 SHFQSCQLALEPQAVSSYLSPGAPLKGRPP-----SPGFQRQRORRR 97

DB 126 QRFPSPCGRSPSR-GSHLRP-----RGSPDPLRRRQHQARQSPAGQRRRRROHR 178

RESULT 2

US-09-252-991A-25176 Application US/09252991A
; Sequence 25176, Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25176
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25176

Query Match      15.7%; Score 84; DB 4; Length 465;
Best Local Similarity 28.0%; Pred. No. 0.31;
Matches 33; Conservative 8; Mismatches 37; Indels 40; Gaps 6;

QY 1 PCR-----PRTAP-----KAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILA 43
Db 46 PGRPARRRSGRRCCAYFTHPGVGTGAPARPAIGCATVAIRPAHPRPSAVGPHAGTLP 105
QY 44 CGFLLGCW----GPSH-----FQSCLOALEPQAVSSYLSGAPLKGRRPPSP 86
Db 106 CG-----WPARSHPGHAGRRGLRDWQSG---GEEVWACSHRRPDAPPKPPPPP 155

RESULT 3
US-09-252-991A-27050
; Sequence 27050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 27050
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-27050

Query Match      15.2%; Score 81; DB 4; Length 245;
Best Local Similarity 27.0%; Pred. No. 0.32;
Matches 24; Conservative 11; Mismatches 48; Indels 6; Gaps 2;

QY 6 RPKAPSHSAPLLGLALLRMHQRHPRARCPPLCVAGILACGFLGCGWSPHFQSCLOALE 65
Db 66 RPPTSAGAMFVPPATTTSTCTCRPAPSCAPFAPSFTACGTRLAAWPPS--ASDC---Q 119
QY 66 PQAVSSYLSGAPLKGRRPSPGFQRROR 94
Db 120 PRAIGRCASRRCCRPASPSPWARRR 148

RESULT 4
US-09-252-991A-30702
; Sequence 30702, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30702
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30702

; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30702

Query Match      15.2%; Score 81; DB 4; Length 314;
Best Local Similarity 30.3%; Pred. No. 0.42;
Matches 33; Conservative 7; Mismatches 35; Indels 34; Gaps 7;

QY 1 PGRPT-RPKAPSHSAPLLGLALLR-MHQRHPRARCPPLCVAGILACGFLGCGWSPSH-FQ 57
Db 114 PGRPPRPAIPQAVPLRPAVRLRPVGHHPRAELP-----GAGGTHRLD 159
QY 58 QSCLOALEPQAVSSYLSGAPLK-----GRPPS-----PGFQR 90
Db 160 RPRRLHYQAAVPGDLA-GIPLEEARLRLSLPENGRCQDQVLYTHPAQR 207

RESULT 5
US-09-252-991A-24664
; Sequence 24664, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24664
; LENGTH: 324
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24664

Query Match      15.1%; Score 80.5; DB 4; Length 324;
Best Local Similarity 28.7%; Pred. No. 0.5;
Matches 33; Conservative 6; Mismatches 43; Indels 33; Gaps 5;

QY 3 RPTRPKAPSHSAPLLGLALLRMHQRH-----RARCPLCVAGILAC-G 45
Db 17 RPRRPSPSPRG-----RHPTTPRAGTGVPRPGWRRARCGR--NAASRACAG 63
QY 46 FLGCGWSPHFQSCLOALEPQAVSSYLSGAPLKGRRPPSP---GFQRRQRORR 97
Db 64 FRWSWAPPEDRDVCLSPASGSRPWWGRRGPPPTPEPPRAGVQTPERHRK 118

RESULT 6
US-09-252-991A-19528
; Sequence 19528, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19528
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19528
```

Query Match 14.9%; Score 79.5; DB 4; Length 510;
Best Local Similarity 28.2%; Pred. No. 1.1;
Matches 35; Conservative 11; Mismatches 37; Indels 41; Gaps 7;

QY 2 GRPTRKAPSHSAPLLGLALLRMH-----QRIHRA-----RCPPCLVAGIL 42
DB 112 GRPERORPDH-ADAVHPADLQCHAGAVRHQAGQRQRAVEERRRGIGACRPAARQA- 169
QY 43 ACGFLLGCWGPSPHFQSQCL--QALEP-----QAVSSYLSFGAPLKGKRPSPGFORQ 93
DB 170 -----GPADRRHRVHRQEPARRQPADDSAEYRRFG--LRGRPGQPGGRRLLR 218
QY 94 RQRR 97
DB 219 RRPR 222

RESULT 7
US-09-252-991A-17616
; Sequence 17616, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17616
; LENGTH: 834
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17616

Query Match 14.6%; Score 78; DB 4; Length 834;
Best Local Similarity 22.7%; Pred. No. 2.7;
Matches 30; Conservative 12; Mismatches 28; Indels 62; Gaps 5;

QY 2 GRPTRKAPSHS-----APLLGLALLRMHQR-----HPRARCPPLCVAGILACGFLGC 50
DB 22 GRPHQGEPRSPGRCVGIAGLPRLAHVORLQPLRAHLQGHCGRCAA----- 70
QY 51 WGPSPHFQSQCLQALEPQAVSSYLSPG-----APLKGRRPSP----- 86
DB 71 -----PHAGRSHRAPAGTQCRRGDVAVLVHRDQALRSRPGDPLQRLSV 115
QY 87 -GFQRRQRQR 97
DB 116 GGYORRRASRQ 127

RESULT 8
US-09-252-991A-23651
; Sequence 23651, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 23651
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23651

Query Match 14.5%; Score 77.5; DB 4; Length 273;
Best Local Similarity 28.2%; Pred. No. 0.87;
Matches 31; Conservative 8; Mismatches 46; Indels 25; Gaps 4;

QY 4 PTERKAPSHSA-----PLLGLALLRMHQRHRA-----RCPPCLVAGILACGF 46
DB 14 PARPPGGRNAAGRPGRGPAAGRRRLRCRRPRAAPRAAAATPAAPRRHAARAGG- 72
QY 47 LLGCWGPSPHFQSQCLQALEPQAVSSYLSPGAPLKGKRPSPGFORQ 96
DB 73 --STPRSPRSPGRTGNVPQP-----APGSPAAGRPPPGDRSSASRR 115

RESULT 9
US-09-252-991A-28706
; Sequence 28706, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28706
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28706

Query Match 14.2%; Score 76; DB 4; Length 445;
Best Local Similarity 27.9%; Pred. No. 2.2;
Matches 34; Conservative 8; Mismatches 40; Indels 40; Gaps 5;

QY 3 RPTRKAPSHSAPLLGLALLRMHQR-----PRARCPPLCVAGILAC 44
DB 100 RLARP-AGLHPATSLGLPLPAQRAHLHQAGRNRPQGQLDGRGPRR-----AA 147
QY 45 GFLLGCWGPSPH--PQSCLOALEPQAVSSYLSPGAPLKGKRPSPGFORQ 94
DB 148 EYTGCRGQLHPSLVDPDLLRACRPAHLRSQSDRRPLRRQPPAPGTRWRATAHDPDQR 207
QY 95 QR 96
DB 208 HR 209

RESULT 10
US-09-252-991A-19363
; Sequence 19363, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19363

Query Match
Best Local Similarity 14.1%; Score 75.5; DB 4; Length 330;
Matches 26; Conservative 12; Mismatches 34; Indels 19; Gaps 3;

QY 1 PGRPTRPKAPSHSAPLLGLA-----LLMHQHPHRCPPPLCVAGILACGFLGCGWPSHF 56
Db 250 PGEPRPRSGWRGAGACGPIVLGQRPAAACGGLCAH-----LPQGG 298
QY 57 QOSCLQALEPQ-----AVSSVSLSPGAPLKGRP 83
Db 299 RDSRUSAQDPWPWTCGRSSMKAAASLRGRP 329

RESULT 11
US-09-252-991A-31668
; Sequence 31668, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31668
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31668

Query Match
Best Local Similarity 14.0%; Score 75; DB 4; Length 244;
Matches 28; Conservative 5; Mismatches 33; Indels 32; Gaps 3;

QY 1 PGRPTRPKAPSHSAPLLGLALRM-HQHPRARCPLCVAGILACGFLGCGWPSHFQOS 59
Db 18 PGGPRRCACRSEATWATLFRHCHPDGRTGCG----- 52
QY 60 CLQALEPQAVSSVSLSPGAPLKGRPPSPGFORQORR 97
Db 53 --QLLDPAVARQSRDPHGG---PGSQRRFSRR 84

RESULT 12
US-09-252-991A-17078
; Sequence 17078, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17078
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17078

Query Match
Best Local Similarity 13.9%; Score 74; DB 4; Length 271;
Matches 31; Conservative 7; Mismatches 39; Indels 22; Gaps 5;

QY 1 PGRPTRPKAPSHSAPLLGLALRMHQHPHRCPPPLCVAGILACGFLGCGWPSHFQ 57
Db 68 PARRLPR-PQRHPHRSLLRRPAGQRRAPQRSAGSAVA--ACGRATGIPRRGGPGHGR 124
QY 58 QSCLEPQAVSSVSLSPGAPLKGRPPSPGFORQORR 96
Db 125 RPSLQ-----PRQG-PRTGPOARRRR 147

RESULT 13
US-09-252-991A-25602
; Sequence 25602, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25602
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25602

Query Match
Best Local Similarity 13.9%; Score 74; DB 4; Length 366;
Matches 33; Conservative 6; Mismatches 33; Indels 36; Gaps 6;

QY 1 PGRPTRPKAPS-HSAPLLGLALRMHQHPHRCPPPLCVAGILACGFLGCGWPSHFQOS 59
Db 276 PGGPPSPPPFPVADGNGWRLR-----RASCPP-----GC-----RS 310
QY 60 CLQALEPQAVSSVSL-----SPGAPLKGRPPS-----PGFORQORR 97
Db 311 SRSRPRVAPSPAREPAGREGGWLGRPPGAPVRRPGRCRDPVRR 358

RESULT 14
US-09-252-991A-29365
; Sequence 29365, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29365
; LENGTH: 213
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29365
```

Query Match 13.8%; Score 73.5; DB 4; Length 213;
Best Local Similarity 28.7%; Pred. No. 1.8;
Matches 31; Conservative 9; Mismatches 39; Indels 29; Gaps 6;

Qy 1 PCRTPPKAPSHSAPL-----LGLALLRMHQHPRARCPCPLCVAGTIA--CGFLLGCWG 52
Db 21 PGVPPGSRPPDHPRPDPFAHLARRVRSRRPAPARCRP---AGILASLCG----- 70

Qy 53 PSHFQSCUQALE-POAVSYLSFGAPLKG--RPPSPGFQORQORR 97
Db 71 -----IRGLAVPGARRRAGLGFMSGLRRPRDQGRRRVFRQWR 110

RESULT 15
US-09-252-991A-29952
; Sequence 29952, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29952
; LENGTH: 286
; TYPE: PXT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29952

Query Match 13.7%; Score 73; DB 4; Length 286;
Best Local Similarity 32.9%; Pred. No. 2.8;
Matches 27; Conservative 11; Mismatches 28; Indels 16; Gaps 5;

Qy 4 PTRPKAPS--HSAPLLGLALLRMHQHHP--RARCPP--LCVAGILACGFLGCGWGPSHFQ 57
Db 212 PRRPASPSAGNAAP---ALPRQHRHPLQLQRRYPPGRCVCSAAVRAAGH-----RQAGQ 261

Qy 58 QSCUQALEPOAVSYLSFGAPL 79
Db 262 RLTLHQEGQGPFCVAPTSPV 283

Search completed: March 13, 2004, 07:46:08
Job time : 5.09219 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:33:19 ; Search time 7.46771 Seconds
(without alignments)

2756.533 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311

Perfect score: 1141

Sequence: 1 AAGGILHLELLVAVGPDVFQ.....QLLSLSAGRACVWDPPRP 214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	297.5	26.1	951	2 T00017	gene ADAMTS-1 prot
2	295	25.9	2165	2 T21371	hypothetical prote
3	284	24.9	837	2 T00355	hypothetical prote
4	272	23.8	1205	2 T18517	procollagen N-endo
5	174.5	15.3	860	2 T16892	hypothetical prote
6	168.5	14.8	1444	2 T18856	angiogenesis inhib
7	152.5	13.4	789	2 S28259	androgen-regulated
8	146.5	12.8	776	2 S28258	androgen-regulated
9	146	12.8	508	2 T22836	hypothetical prote
10	135.5	11.9	429	2 A42972	coagulation factor
11	135.5	11.9	617	2 S48160	metalloproteinase
12	135.5	11.9	814	2 G02390	disintegrin-like m
13	133.5	11.7	201	1 HytV2	trimerelysin II (E
14	127	11.1	600	2 I49281	fertilin alpha pre
15	124	10.9	202	1 HVRSR	ruberylsin (EC 3.4
16	123	10.8	203	2 A59421	metalloproteinase
17	123	10.8	414	1 HYRAC	atrolysin C (EC 3.
18	123	10.8	414	2 S41609	atrolysin C (EC 3.
19	121.5	10.6	478	2 JC4880	fibrinolytic metal
20	121.5	10.6	957	2 T15976	hypothetical prote
21	120	10.5	616	2 A55796	ecarin precursor -
22	118	10.3	480	1 A30085	trigramin precurs
23	117	10.3	411	1 HYSNFA	fibrolase (EC 3.4.
24	115	10.1	481	2 JC4342	fibrinolytic prote
25	115	10.1	903	2 S60257	meltrin alpha - mo
26	112.5	9.9	202	1 HytVH2	hemorrhagic protei
27	112.5	9.9	481	2 S43125	trimucin precursor
28	112	9.8	419	2 A59414	metalloproteinase
29	110.5	9.7	826	2 A60385	monocyte surface a

30 109.5 9.6 549 2 S48169 metalloproteinase
31 109.5 9.6 610 2 JC7530 vascular apoptosis
32 109 9.6 419 2 S41607 atrolysin A (EC 3.
33 108 9.5 655 2 JC7850 disintegrin and me
34 108 9.5 660 2 S71949 metalloproteinase
35 107.5 9.4 407 2 S66260 metalloproteinase
36 107 9.4 203 2 S46443 adamalysin (EC 3.4
37 106.5 9.3 478 2 A43296 atrolysin E (EC 3.
38 105 9.2 952 2 T18900 disintegrin and me
39 104.5 9.2 609 2 S55270 catrocollastatin p
40 103.5 9.1 200 2 S15111 hemorrhagic factor
41 103 9.0 414 2 S41608 atrolysin B (EC 3.
42 103 9.0 416 2 A37877 hemorrhagic protei
43 103 9.0 488 2 JC7969 metalloproteinase 1
44 102.5 9.0 478 2 JQ1301 hemorrhagic protei
45 97 8.5 825 2 S55060 fertilin alpha-II

ALIGNMENTS

RESULT 1

T00017

gene ADAMTS-1 protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000

C:Accession: T00017

R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.

Genomics 46, 466-471, 1997

A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 ge

A:Reference number: Z14055; MUID:98110583; PMID:9441751

A:Accession: T00017

A>Status: Preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-951 <XUN>

A:Cross-references: EMBL:AB001735; NID:G2809056; P:DN:BAA24501.1; PID:G2809057

A:Experimental source: strain 129SVJ

C:Genetics:

A:Gene: ADAMTS-1

A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2

C:Superfamily: thrombospondin type 1 repeat homology

F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match	26.1%	Score	297.5;	DB	2;	Length	951;
Best Local Similarity	29.8%;	Pred. No.	7.2e-13;				
Matches	64;	Conservative	36;	Mismatches	101;	Indels	15;
						Gaps	4;
QY	7	HLELLVAVGPDVFQAHQEDTERVYLTNLNIGAEILLRDPISLGAQFRVHLVQWVILTEPGA	66				
Db	243	YVETMLVADQSMADFHGSGLKHVLLTLFSAARFYKPSIRNSISLVVVKILVIYEEQK	302				
QY	67	PNITANLTSSLVCGNSQINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTQLGG	126				
Db	303	FEVTSNAALTIRNCWQKQNSDRDPEHYDTAILFTRODL-CGSHTCDTLGMADVGT	361				
QY	127	ASPTVSWCLITETGDLGVTHAIEHSGLEHGDGPG-----SCGSGSHWAS-----	177				
Db	362	VCDPFSRSCSVIEDDGLQAFTTAHELGVFNMPDHDKACASLNGVTDGSHLWASMLSSL	421				
QY	178	DGAAPRAGLAWSPCSRRLLSLSAGRACVWDPPR	213				
Db	422	DHSQP-----WSPCSAYMVTSLFDNGHGECLMDKRP	452				

RESULT 2

T21371

hypothetical protein F5H8.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999

C:Accession: T21371; T24896

R:Gajadaty, S.

submitted to the EMBL Data Library, February 1996

A:Reference number: Z19413

A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone F25H8
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <W12>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1, 52/1, 135/2, 193/3, 216/1, 266/1, 495/2, 547/3, 584/3, 634/2, 744/1, 81

Query Match 25.9%; Score 295; DB 2; Length 2165;
Best Local Similarity 30.4%; Pred. No. 3.3e-18;
Matches 70; Conservative 30; Mismatches 104; Indels 26; Gaps 3;
QY 7 HLELLVAGPDVFOAHQEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
DB 281 YVEVLVADTKMYEYHGRSLEDYVTLFSTVASIYHQSLRASINVVVVKLIVLKTENAG 340
QY 67 FNITANLTSSLLSVCGWSQINPEDDTPDGHADLVLYITRFDLELPGNROVRGVTLGG 126
DB 341 PRITQAQTLQDFCRQYVNPDDSSVOHVDVAILLTRKIDCRSQKCDTLGLAELGT 400
QY 127 ACSPTWSCLITDGTDLGVITIAHEIGHSGFLEHGDGAPGSCGP----- 170
DB 401 MCDMQSKALIEDNGLSAAFTIAHELGHVFSIPHDDE--RKCTYMPVNVKVCYKSTQKTFD 458
QY 171 -----SGHVMSDGAAPRAGLAWSPCSRQLLSLL--SAGRACVWDPP 212
DB 459 KTQFQNFHIMPTLEYNTHPWSWSPCSAGMLERFLENRGOTQCLFDQ 508

RESULT 3
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:9840380; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <1SH>
A:Cross-references: EMBL:AB014588; NID:G3327189; PIDN:BAA31663.1; PID:G3327190
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
A:Superfamily: thrombospondin type 1 repeat homology
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 24.9%; Score 284; DB 2; Length 837;
Best Local Similarity 32.5%; Pred. No. 1e-17;
Matches 69; Conservative 27; Mismatches 106; Indels 10; Gaps 4;
QY 8 LELLVAGPDVFOAHQEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 67
DB 220 VETLVVADDKMAAFHAGLKYLLTVMAAAKAFKHPRIKPNVSLVTVTLVILSGSEGP 279
QY 68 NITANLTSSLLSVCGWSQINPEDDTPDGHADLVLYITRFDLELPGNROVRGVTLGGA 127
DB 280 QVGPAAQTLSFCAWQKGLNTPEDSDPDHFTAILFTQDL-CGVSTCDTLGMADVGVTV 338

QY 128 CSPTWSCLITDGTDLGVITIAHEIGHSGFLEHGDGAPGSC-----GP----SGHVMSDGA 180
DB 339 CDPARSCALVEDDGLQSAFTAHELGHVFNMLHDNS--KPCISLNGSLSTRHVPWMA 396
QY 181 APRAGLAWSPCSRRLLSLSAGRACVWDPP 212
DB 397 HVDPEEPWSPCSARFITDPLDNGYGHCLLDKP 428

RESULT 4
T18517
procollagen N-endorpeptidase (EC 3.4.24.14) I - bovine
N:Alternate names: procollagen N-proteinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18517
R:Collige, A.; Nusgens, B.V.; Lapiere, C.M.
submitted to the EMBL Data Library, February 1996
A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A:Reference number: Z18941
A:Accession: T18517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1205 <COL>
A:Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1
A:Experimental source: skin
C:Genetics:
A:Gene: PC I-NP
C:Function:
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior t
C:Keywords: hydrolase; metalloproteinase

Query Match 23.8%; Score 272; DB 2; Length 1205;
Best Local Similarity 33.9%; Pred. No. 2e-16;
Matches 74; Conservative 36; Mismatches 94; Indels 14; Gaps 8;
QY 1 AAGGILHLELLVAGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVI 59
DB 255 AADDYNIIEVLGVDDSVVQFHGTEHVQKYLTLNMVINEYHDSLGAHNVLVRLIL 314
QY 60 LTEPEGAPNI-TANLTSSLLSVCGWSQ-TINPEDDTPDGHADLVLYITRFDLELPGNRQ 117
DB 315 LSYGKMSLIEIGNPQSLENVCRWAYLQKPTDDEYH-DHAIFLTRQDF---GPSG 369
QY 118 VRGVTLQAGACPTWSCLITDGTDLGVITIAHEIGHSGFLEHGDGAPGSCGCPs---GHV 174
DB 370 MQGYAVTGMCHPVRSCITLNHEDGFSSAFVVAHETGTVLGMEDHG-QGNRCGDEVRLGSI 428
QY 175 MASDGAAPRAGLAWSPCSRRLLSLSAGRACVWDPP 212
DB 429 MAPLVQAAPHRFHWRSRCSQELSRYLHS--YDCLRDDP 464

RESULT 5
T16892
hypothetical protein T19D2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16892
R:Bentley, D.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid T19D2.
A:Reference number: Z18599
A:Accession: T16892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-860 <BEN>
A:Cross-references: EMBL:U42846; NID:g1125809; PID:g1125810; PIDN:AAA83600.1; CESP:T191
C:Genetics:
A:Gene: CESP:T19D2.1
A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 5

Query Match 15.3%; Score 174.5; DB 2; Length 860;
Best Local Similarity 24.3%; Pred. No. 9.7e-08;
Matches 54; Conservative 35; Mismatches 87; Indels 45; Gaps 6;

QY 5 ILHLELLVAVGPDVF-----QAHQEDTERYVLTNINIGAEILLRDPISLGAQFRVHLVK 56
DB 132 IITVELAVFADDAAMDFKQYKGAABENMHTFIMAVNNIDVLTORLLQPRINIKIVR 191
QY 57 MYLITEPEGAPNI-----TANLTSSLVCGWSQINPEDDTPDGHADLVLYITRFDL 109
DB 192 YEIL-----KNIPHLMNARKSGNDVRLDLDAFCQYQNEINPNPDADPRHWDHALLFGSYDL 248
QY 110 ELPDGNRQVRGVTQLGGACSPWTSCLITEITDGFOLGVTIAEIGHSGFLEHGDG----- 162
DB 249 H-RNGVKTVAGVAPVKMGCSGVTSCTINEGLDFGSVFVWTHMGHSLGMYHDGNECDLR 307
QY 163 -----APGSCGPGSHVMASDGAAPRAGLAWSPCSRRLSLL 200
DB 308 CCIMSPSVGSGKT-----HWSQCSVNEWATEV 334

RESULT 6
T18856
angio genesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18856; T24653
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: T18856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL>
A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone C02B4
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19917
A:Accession: T24653
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <W12>
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1
A:Experimental source: clone T07C5
C:Genetics:
A:Gene: CESP:C02B4.1
A:Map position: X 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 566

Query Match 14.8%; Score 168.5; DB 2; Length 1444;
Best Local Similarity 26.4%; Pred. No. 6.4e-07;
Matches 55; Conservative 30; Mismatches 98; Indels 25; Gaps 6;

QY 8 LELLVAVGPDVFOAQHEDTERYVLTNINIGAEILLRDPISLGAQFRVHLVYKLVILTEPEGAP 67
DB 245 LEFSLALINNHHVLYQDTPNLNDIVIVRYEMWRTQPSALSTGVHK----- 291
QY 68 NITANLTSSLVCGWSQINPEDD-TDPGHADLVLYITRFDLDPGNRQVRGVTQLGG 126
DB 292 --NGQAQSLDLAFCRYQAQHNWPGTDLTMNHYDGVLLTYGDIY--HTTTSVAGVAPVAR 347
QY 127 ACSPTWSCSLITEITDGFOLGVTIAEIGHSGFLEHGDGPGSCGPGSHVMASDGAAPRAGL 186
DB 348 MCDPLFACSLVEGLHLGRSFLAHMGHNNWGHVDGVQNO-CNKGCCLLMSAVNGAGKT-- 404
QY 187 AWSPCSRQ-----LLSLLSAGRARCVD 210
DB 405 TWSDCSVREFNAFLQLDSESGRGNCLRD 432

RESULT 7
S28259

Query Match 12.8%; Score 146.5; DB 2; Length 776;
Best Local Similarity 26.0%; Pred. No. 3e-05;
Matches 60; Conservative 35; Mismatches 79; Indels 57; Gaps 12;

QY 7 HLELLVAVGPDVFPQ--AQEDTER-----YVLTNINIGAEILLRDPISLGAQFRVH 53
DB 200 YIELFIVADDTVVRNSHPNKLNRIRWGVMVFNMYIKTLNI-----HVT 245
QY 54 LVQNVILTEPEGAPNITANLTSSLVCGWSQINPEDDTPDGHADLVLYITRFDLE---- 110
DB 246 LVGIEIWTW-EDKIELHSNITETLLRFSWQERI-----LTKRDFDHVV 289
QY 111 LPDGN---ROYRGVTQLGGACSPWTSCLITEITDGFOLGV---TTAEIGHSGFLEHGDGAP 164
DB 290 LLSGRWYTHVQGISYPAGMCLPYSTSIINKLLPDNIIANRMAHQHNLGMQHDPEP 349

Query Match 13.4%; Score 152.5; DB 2; Length 789;
Best Local Similarity 29.1%; Pred. No. 8.8e-06;
Matches 50; Conservative 27; Mismatches 68; Indels 27; Gaps 9;

QY 51 RVHLVKNVILTEPEGAPNITANLTSSLVCGWSQIT-NPEDDTPDGHADLV-----LYIT 105
DB 243 RVTLTGWEIWSAGDEI-EIVSNLESTLLHSTWQETVLKKEKDFD--HVILLSGKWLYTS 299
QY 106 RPDELPDGNRQVRGVTQLGGACSPWTSCLITEITDGF-----LGVTTIAEIGHSGFLEHGDG 162
DB 300 -----MQGIAYPGGICQTLRSCSVKOLLPDVNIIGNRMAHQHLSGLGMRHDD 347
QY 163 APGSCGPGSHVMASDGAAPRAGLAWSPCSRRLSLLSAGRARCVDWPPRP 214
DB 348 FPCT-C-PLGKCYMGAGSIP--AIKFSKCSQTYQQFLKNQKPAICILNNPLP 395

RESULT 8
S28258
androgen-regulated epididymal protein precursor - crab-eating macaque
C:Species: Macaca fascicularis (crab-eating macaque)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jul-2000
C:Accession: S28258
R:Perry, A.C.F.; Jones, R.; Barker, P.J.; Hall, L.
Biochem. J. 286, 671-675, 1992
A:Title: A mammalian epididymal protein with remarkable sequence similarity to snake v
A:Reference number: S28258; MUID:93038491; PMID:1417724
A:Accession: S28258
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-776 <PER>
A:Cross-references: EMBL:X66139; NID:G38062; PIDN:CAA46929.1; PID:G38063
C:Superfamily: mouse meltrin alpha; disintegrin homology
F:1-25/Domain: signal sequence #status predicted <SIG>
F:26-776/Product: androgen-regulated epididymal protein #status predicted <MAT>
F:336/Region: defective catalytic site
F:401-483/Domain: disintegrin homology <DIS>

QY 165 GSGCGPSGH-VMASDGAAPAGLAWSPCRRQLLSLSAGRARCVCWDPERP 214
DB 350 CT-C-PSGKCVMSDGSIP--ALKFSKCSNQHQYLDKDYKPTCMLNIPFP 396

RESULT 9
T22836
hypothetical protein F57B7.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22836
R:Lennard, N.
submitted to the EMBL Data Library, June 1996
A:Reference number: Z19623
A:Accession: T22836
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-508 <WIL>
A:Cross-references: EMBL:Z74037; PIDN:CAA98493.1; GSPDB:GNC00023; CESP:F57B7.4
A:Experimental source: clone F57B7
C:Genetics:
A:Gene: CBSP:F57B7.4
A:Map position: 5
A:Introns: 45/3; 137/2; 221/2; 256/2; 306/2; 409/3; 451/3

Query Match 12.8%; Score 146; DB 2; Length 508;
Best Local Similarity 31.1%; Pred. No. 2e-05;
Matches 55; Conservative 24; Mismatches 86; Indels 12; Gaps 7;

QY 43 DPSLGAQFRVHLVKMWLTPEEGAPNITANTSSLLSCVGSQ-TINPEDDTD-PGHADL 100
DB 192 DALISSDMPKKLRKFVDITLLEMOENNSTEMTLKIDSKKAIKDKFTILWKEQGLPRHEHA 251

QY 101 VLITTRDELPLDGNQVRGVQTLGGACSPWSCLTEDTGFDL-GVITIAHIGHSFGL 159
DB 252 VL-ITKFDLLISNGSATQGVAYGVNIGENGSSVVEDIGAGLISLMAHIGHSGLAL 310

QY 160 HDGA-PGSGCGPSGH---VMASDGAAPAGL-----AWSPCRRQLLSLSAGRARCVC 208
DB 311 HDGAYETACDSNDNYLMAVAVSGSADRQSFNSRMENCSINSIENLKPTANCV 367

RESULT 10
A42972
coagulation factor X activating enzyme (EC 3.4.24.-) heavy chain - Russell's viper
N:Alternate names: RVV-X-heavy chain
C:Species: Vipera russelli (Russell's viper)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 19-May-2000
C:Accession: A42972
R:Takeya, H.; Nishida, S.; Miyata, T.; Kawada, S.; Saisaka, Y.; Morita, T.; Iwanaga, S.
J. Biol. Chem. 267, 14105-14117, 1992
A:Title: Coagulation factor X activating enzyme from Russell's viper venom (RVV-X). A nd
A:Reference number: A42972; MUID:9232516; PMID:1629211
A:Contents: V. r. siamensis
A:Accession: A42972
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-429 <TAK>
A:Experimental source: venom
A:Note: sequence extracted from NCBI backbone (NCBIP:108405)
C:Comment: This metalloproteinase has disintegrin (platelet aggregation inhibitor)-like
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: hydrolase; metalloproteinase; venom; zinc
F:212-296/Domain: disintegrin homology <DIS>
F:146/Active site: Glu #status predicted

Query Match 11.9%; Score 135.5; DB 2; Length 429;
Best Local Similarity 25.0%; Pred. No. 0.00015;
Matches 53; Conservative 30; Mismatches 76; Indels 53; Gaps 9;

QY 45 SLGAQFRVHLVKMWLTPEEGAP--NITA-----NLTSLLSCVGV 83
DB 3 STSAQFNKIFELVIIVDHSMAKKCNSTATNTKIYBIVNSANBIFNPLNTHVLIGVEFW 62

QY 84 -----SQTINPEDD-----TDPGHADLVLI-TRFDLELPDGNQVRGVQ 123
DB 63 CDRLINTSSADETLNSFGWFRASDLMTKSHDNALLFTDMRFDL-----NTLGITF 115

QY 124 LGGACSPWSCLTEDTG---FDLGVITIAHIGHSFGLHDGAPSGCGPSGHVNASDGA 180
DB 116 LAGMCOAYRSVELVQEGNRNFKTAVIMAHLSHNLGMYHDG-KNCICNDSSCYM-SPVL 173

QY 181 APAGLAWSPCRRQLLSLSAGRARCVCWDP 212
DB 174 SDQPSKLFNSCSITHDYQRYLTRYKPKCIFNPP 205

RESULT 11

S48160
metalloproteinase (EC 3.4.24.-) H-I precursor - carpet viper
N:Contains: disintegrin
C:Species: Echis pyramidum leakeyi
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jul-2000
C:Accession: S48160
R:Paine, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.
Eur. J. Biochem. 224, 483-488, 1994
A:Title: Cloning of metalloproteinase genes in the carpet viper (Echis pyramidum leakeyi)
A:Reference number: S48160; MUID:95010025; PMID:7925363
A:Accession: S48160
A:Molecule type: mRNA
A:Residues: 1-617 <PAI>
A:Cross-references: GB:X78970; NID:G763092; PIDN:CAA55565.1; PID:G763093
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: hydrolase; metalloproteinase; venom
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-192/Domain: propeptide #status predicted <PRO>
F:193-617/Product: metalloproteinase H-I #status predicted <MAT>
F:403-485/Domain: disintegrin homology <DIS>
F:337/Active site: Glu #status predicted

Query Match 11.9%; Score 135.5; DB 2; Length 617;
Best Local Similarity 28.6%; Pred. No. 0.00023;
Matches 30; Conservative 24; Mismatches 38; Indels 13; Gaps 3;

QY 115 NRQVRGVQTLGGACSPWSCLTED---TGFDLGVITIAHIGHSFGLSHD---GAPGSG 167
DB 298 NGOTLGARMRGMCSPNSGVITQDYKNYLLVITMAHELGNLGMHDHNGNCNCPDTS 357

QY 168 CGPSGHVWASDGAAPAGLAWSPCRRQLLSLSAGRARCVCWDP 212
DB 358 C-----IMSAVAGPEPVFSFNSCNRNDYFRNSDQSKCIDNKP 396

RESULT 12

G02390
disintegrin-like metalloproteinase MDC15 (EC 3.4.24.-) - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 31-Dec-2000
C:Accession: G02390; PC4263
R:Herren, B.; Raines, E.W.; Ross, R.
submitted to the EMBL Data Library, January 1996
A:Reference number: H01157
A:Accession: G02390
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-814 <HER>
A:Cross-references: EMBL:U46005; NID:G1335871; PIDN:AACS1112.1; PID:G1335872
Biochem. Biophys. Res. Commun. 230, 335-339, 1997
A:Title: Expression of members of a novel membrane linked metalloproteinase family (AD)

A:Reference number: PC4263; MUID:97168971; PMID:9016778
A:Accession: PC4263
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-461 <MCK>
A:Experimental source: articular chondrocyte

C:Comment: This protein is a membrane bound protein and involved in cell/cell and cell/m
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: hydrolase; metalloprotease; zinc
F:420-503/Domain: disintegrin homology <DIS>
F:348,352,358/Binding site: zinc (His) #status predicted
F:1349/Active site: Glu #status predicted

Query Match 11.9%; Score 135.5; DB 2; Length 814;
Best Local Similarity 32.0%; Pred. No. 0.00032;
Matches 54; Conservative 22; Mismatches 62; Indels 31; Gaps 10;

QY 72 NUTSSLLSVCGWSQTINPEDTPG-HADLVLYITRFDL--ELPDGNGRQVGTOLGG-- 126

Db 254 NVRVALVGLAEATQRDVLEISPAVLEFLHWRRAHLLPRLPDSDAQLVTGTSFSGPT 313

QY 127 -----ACSPWSSCLITDTPD-LGV--TIAHEICHSEGLEHGDGAPGSGC---GRSG 172

Db 314 VGMIAQNSICSPDFSGVNDHSTISLGVASSIAHEHLSGLDHD-LFGNSCPGCGPAP 372

QY 173 H---WASDGAAPRAGLAWPCSRRLSLLSAGRARCVDW----PP 212

Db 373 AKTCINEASTDFLP--GLNFSNCSRRALEKALLDGMSCFLERLPSP 419

RESULT 13

HYTV2

trimerelysin II (EC 3.4.24.53) - habu

N:Alternate names: proteinase H2; Trimeresurus metalloendopeptidase II

C:Species: Trimeresurus flavoviridis (habu)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Aug-1994

C:Accession: J00037

R:Takeya, H.; Arakawa, M.; Miyata, T.; Iwanaga, S.; Omori-Sato, T.

J. Biochem. 106, 151-157, 1989

A:Title: Primary structure of H-2-proteinase, a non-hemorrhagic metalloproteinase, isolated

A:Reference number: J00037; MUID:89380121; PMID:2777746

A:Accession: J00037

A:Molecule type: protein

A:Residues: 1-201 <FAK>

C:Comment: This proteinase is a major venom non-hemorrhagic metalloproteinase.

C:Superfamily: atrolysin C

C:Keywords: hydrolase; metalloprotease; pyroglutamic acid; venom; zinc

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:117-196,158-180,160-163/disulfide bonds: #status experimental

F:142,146,152/Binding site: zinc (His) #status predicted

F:143/Active site: Glu #status predicted

Query Match 11.7%; Score 133.5; DB 1; Length 201;

Best Local Similarity 21.2%; Pred. No. 9e-05; DB 1; Length 201;

Matches 46; Conservative 48; Mismatches 90; Indels 33; Gaps 9;

QY 7 HLELLVAVGPDVFOAQHEDTER-----YVLTNLNIGAEILLRDPGLGAQF-RVHLVVKVI 59

Db 7 YIELAVVDHGMVKYKYNQSDKIKVRVHQVYNHIN---EMVRPLNIAISLNLQIWSKD 63

QY 60 LTFEGAPNITANLTSSLLSVCGWSQTI-NPEDDTPGHADLVLYITRFDLELPDGNRQV 118

Db 64 LITVKSASNT-----LESFGNWRKTVLLQKQNDCAH-----LLTATNL-----NDNT 107

QY 119 RGVITQGGACSPWSSCLITDTPG---FDLGVITIAHEICHSEGLEHGDGAPGSGCGPSGHVM 175

Db 108 IGLAYKGMCKPKLSVGLVDYSPNVFMVAVTMTHELGNLGMEDDKCKKCEA---CI 164

QY 176 ASDGAAPRAGLAWPCSRRLSLLSAGRARCVDPP 212

Db 165 MSDVISDKPSKLPSCDKNDYQTFLTKNYPOCILNAP 201

RESULT 14

I49281

fertilin alpha precursor - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jun-2000

C:Accession: I49281

R:Wolfsberg, T.G.; Straight, P.D.; Gerena, R.L.; Huovila, A.P.; Primakoff, P.; Myles, I
Dev. Biol. 169, 378-383, 1995

A:Title: ADAM, a widely distributed and developmentally regulated gene family encoding
A:Reference number: I48100; MUID:95269891; PMID:7750654

A:Accession: I49281

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-600 <RES>

A:Cross-references: EMBL:U22056; NID:g965009; PIDN:AAA7920.1; PID:g965010

C:Genetics:

A:Gene: ADAM 1

C:Superfamily: unassigned disintegrins; disintegrin homology

F:246-326/Domain: disintegrin homology <DIS>

F:180/Active site: Glu #status predicted

Query Match 11.1%; Score 127; DB 2; Length 600;

Best Local Similarity 26.8%; Pred. No. 0.0013;

Matches 57; Conservative 32; Mismatches 97; Indels 28; Gaps 10;

QY 7 HLELLVAVGPDVFOAQHEDTERVYVLTNLNIGAEILLRDPGLGAQFVRHLVVKVILTEPGA 66

Db 45 YVEMFVVVNHQRQFMGNSINETVQAVNDIIA-LANSFTRGINTEVVLVGLIEWT--EGD 101

QY 67 P-NITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDL--ELPDGNGR--QVRGVT 122

Db 102 PIEVPVDLQTLLENFNRQ-----EKLVRVRVHDVAHLVGHRRPGENEGA 148

QY 123 QLGACSPWSSCLUIT-----EDTGFDLGVITIAHEICHSEGLEHGDGAPGSGCGSHVMASD 178

Db 149 FLRGACSGFEAAAVEAPHEDV-LLFAALMAHELGNLGIQHD-HPTCTCGPKHFCLMGE 206

QY 179 GAAPRAGLAWPCSRRLSLLSAGRARCVDPP 212

Db 207 KIGKDSG--FSNCSDFLFLHHRGACLLDEP 238

RESULT 15

HYRSR

ruberlysin (EC 3.4.24.48) - red diamond rattlesnake

N:Alternate names: Crotalus ruber metalloendopeptidase II; hemorrhagic metalloproteinase

C:Species: Crotalus ruber (red diamond rattlesnake)

C:Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 05-Aug-1994

C:Accession: JX0139

R:Takeya, H.; Onikura, A.; Nikai, T.; Sugihara, H.; Iwanaga, S.

J. Biochem. 108, 711-719, 1990

A:Title: Primary structure of a hemorrhagic metalloproteinase, HT-2, isolated from the

A:Reference number: JX0139; MUID:91185324; PMID:2081731

A:Accession: JX0139

A:Molecule type: protein

A:Residues: 1-202 <FAK>

C:Superfamily: atrolysin C

C:Keywords: hydrolase; metalloprotease; pyroglutamic acid; venom; zinc

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:117-197,157-164/disulfide bonds: #status experimental

F:142,146,152/Binding site: zinc (His) #status predicted

F:143/Active site: Glu #status predicted

Query Match 10.9%; Score 124; DB 1; Length 202;

Best Local Similarity 23.4%; Pred. No. 0.00066;

Matches 50; Conservative 38; Mismatches 100; Indels 26; Gaps 9;

QY 7 HLELLVAVGPDVFOAQHEDTERVYVLTNLNIGAEILLRDPGLGAQFVRHLVVKVILTEPGA 66

Db 7 YIELVVVDHHRMPKYNKSLNT-IRTVHEIVNFINFEYRSINIRVSLTDLIWSDDFI 65

QY 67 PNITANLTSSLLSVCGWSQTI---NPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTQ 123

Db 66 -TVQSSAKNTLHSGFEWKRKSVLLNRKRD---NAQLLTAIVLDDYTL-----GLAY 112

QY 124 LGGACSPWSSCLITDTPG---TGFDLGVITIAHEICHSEGLEHGDGAPGSGC---GPSCHVMASD 178

Db 113 LNSMCHPRNSVGLIQDHPISNLLMGVTMAHELGNLGMEDH---GKDCLRGASLCIM-RP 168

Qy 179 GAAPRAGLAWPCSRROLLSLLSAGRARCVDPP 212
Db 169 GLTPGRSYEFSDASMRYYQKFLDQYKPCILNKP 202

Search completed: March 13, 2004, 07:44:35
Job time : 8.46771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 4.45833 Seconds
(without alignments)
2499.367 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311

Perfect score: 1141

Sequence: 1 AAGGILHLELLVAGPDVFQ.....QLLSLSAGRCVWDPPRP 214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	361	31.6	997	1 ATS7_HUMAN	Q9ukp4 homo sapien
2	336	29.4	1593	1 A12_HUMAN	P58397 homo sapien
3	330	28.9	1911	1 A20_HUMAN	P59510 homo sapien
4	316.5	27.7	1906	1 A20_MOUSE	P59511 mus musculus
5	307.5	27.0	1224	1 A16_HUMAN	Q8te57 homo sapien
6	304.5	26.7	967	1 A16_HUMAN	Q9uh18 homo sapien
7	302.5	26.5	1077	1 A10_HUMAN	Q9h324 homo sapien
8	299.5	26.2	967	1 A11_RAT	Q9wug1 rattus norv
9	297.5	26.1	968	1 A11_MOUSE	P97857 mus musculus
10	292	25.6	1935	1 A19_HUMAN	Q9p2n4 homo sapien
11	285	25.0	890	1 A18_HUMAN	Q9up79 homo sapien
12	284	24.9	837	1 A18_HUMAN	Q75173 homo sapien
13	281	24.6	630	1 A18_RAT	Q9esp7 rattus norv
14	275	24.1	1213	1 A18_MOUSE	Q8c9w3 m adamts-2
15	274	24.0	1095	1 A17_HUMAN	Q8te56 homo sapien
16	272	23.8	1205	1 A17_BOVIN	P79331 b adamts-2
17	271.5	23.8	950	1 A15_HUMAN	Q8te58 homo sapien
18	271	23.8	1211	1 A15_HUMAN	Q95450 h adamts-2
19	269.5	23.6	1223	1 A14_HUMAN	Q8wx88 homo sapien
20	261	22.9	1081	1 A18_HUMAN	Q8te60 homo sapien
21	260	22.8	1205	1 A18_MOUSE	O15072 homo sapien
22	256	22.4	905	1 A18_MOUSE	P57110 mus musculus
23	254.5	22.3	930	1 A18_MOUSE	Q9r001 mus musculus
24	249	21.8	860	1 A18_HUMAN	Q9ukp5 homo sapien
25	245.5	21.5	930	1 A18_HUMAN	Q9una0 homo sapien
26	240	21.0	1207	1 A19_HUMAN	Q8te59 homo sapien
27	236	20.7	1210	1 A19_MOUSE	P59509 mus musculus
28	166	14.5	788	1 A17_MOUSE	O35227 mus musculus
29	160	14.0	864	1 A15_MOUSE	O88839 mus musculus
30	157.5	13.8	245	1 A15_BOVIN	Q9tt593 bos taurus
31	157	13.8	816	1 A15_RAT	Q9gyv0 r adam 15 p
32	152.5	13.4	789	1 A17_RAT	Q63180 rattus norv
33	150	13.1	754	1 A17_HUMAN	Q9h2u9 homo sapien

ALIGNMENTS

RESULT 1

ID	ATS7_HUMAN	STANDARD;	PRT;	997 AA.
AC	Q9UKP4;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).			
GN	ADAMTS7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99395124; PubMed=10464288;			
RA	Hurskainen T.L., Hirohata S., Seldin M.F., Apte S.S.;			
RT	"ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases."			
RL	J. Biol. Chem. 274:25555-25563 (1999).			
CC	-1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).			
CC	-1- TISSUE SPECIFICITY: Expressed in heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas.			
CC	-1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.			
CC	-1- PM: The precursor is cleaved by a furin endopeptidase (By similarity).			
CC	-1- SIMILARITY: Belongs to peptidase family M12B.			
CC	-1- SIMILARITY: Contains 1 disintegrin-like domain.			
CC	-1- SIMILARITY: Contains 2 TSP type-1 domains.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).			
CC	EMBL; AF140675; AAD56358.1; -			
DR	HSSP; P15167; 1ATL.			
DR	MEROPS; M12.231; -			
DR	Genew; HGNC:223; ADAMTS7.			
DR	MM; 605009; -			
DR	GO; GO:0008237; F:metallopeptidase activity; TAS.			
DR	InterPro; IPR001762; Disintegrin.			
DR	InterPro; IPR001818; Pept_M10A_M12B.			
DR	InterPro; IPR006025; Pept_M_Zn_BS.			
DR	InterPro; IPR001590; Peptidase_M12B.			
DR	InterPro; IPR002870; Peptidase_M12B_N.			
DR	InterPro; IPR000884; TSP1.			
DR	InterPro; IPR008085; TSP 1.			
DR	Pfam; PF01562; Pep_M12B_propep; 1.			

Q28475 macaca fasc
Q13444 homo sapien
P70505 rattus norv
P20165 trimeresuru
Q9b211 homo sapien
Q35674 mus musculu
Q9r160 mus musculu
Q60813 mus musculu
Q9r158 mus musculu
Q8r134 mus musculu
P20897 crocalus ru
P18619 trimeresuru

34 146.5 12.8 776 1 AD07 MACPA
35 135.5 11.9 814 1 AD15 HUMAN
36 134 11.7 789 1 AD01 RAT
37 133.5 11.7 201 1 HR2 TRIPL
38 133.5 11.7 813 1 AD33 HUMAN
39 132 11.6 920 1 AD19 MOUSE
40 128 11.2 761 1 AD24 MOUSE
41 127 11.1 791 1 AD1A MOUSE
42 126 11.0 697 1 AD2B MOUSE
43 125 11.0 806 1 AD1B MOUSE
44 124 10.9 202 1 HRT2 CRORU
45 124 10.9 483 1 DISF TRIPL


```
FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 393 393 BY SIMILARITY.
FT METAL 396 396 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1300 1300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1371 1371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1503 1503 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1593 AA; 177545 MW; 0779F48E63BD83A3 CRC64;

Query Match 29.4%; Score 336; DB 1; Length 1593;
Best Local Similarity 34.9%; Pred. No. 8.7e-21;
Matches 74; Conservative 39; Mismatches 91; Indels 8; Gaps 5;

QY 8 LELLVAVGDFVQAH-QEDTERVYLTNLNIGAEULRDPISLGAQFRVHLVKVILTEPEGA 66
Db 248 VETLVVADTKMIEYHGSSENVESILITNNMVTGLFHPNSIGNAHIVVVRLLIEEBEQ 307
QY 67 FNITANLTSSLSVCGWSQTINPDDTDPGHADLVLYITRDELPLDGNR--QVRGVTL 124
Db 308 LKIVHAETLSFCWKQKINPKSDLPVHVDVAVLLTRKDI-CAGFNRCETLGLSHL 366
QY 125 GACSTPTWCLINTEGDFDLGVITIAHIGHSFLEHGAPOGSGCPG---HVMASDGA 181
Db 367 SGWCQPHRSNINEDSLPLAFTIAHELHSGFGIHDGKE-NDCEPVGRHPYIMSRQLQY 425
QY 182 PRAGLAWSCSRQLLSLSAGRCARCVDPR 213
Db 426 DPTPLTWSKCEBYITRFLDRGMGFCCLDTPK 457

RESULT 3
ID AT20 HUMAN STANDARD; PRT; 1911 AA.
AC P59510;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
DE TS20).
GN ADAMTS20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wight T.N., Leduc R., Apte S.S.;
RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RT subfamily related to Caenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513 (2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Liver;
RX MEDLINE=22566039; PubMed=12562771;
RA Ilamazares M., Cal S., Quesada V., Lopez-Otin C.;
RT "Identification and characterization of ADAMTS-20 defines a novel
RT subfamily of metalloproteinases-disintegrins with multiple
RT thrombospondin-1 repeats and a unique GON domain.";
RL J. Biol. Chem. 278:13382-13389 (2003).
```

```
CC -!- FUNCTION: May play a role in tissue-remodeling process occurring
CC in both normal and pathological conditions.
CC -!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59510-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59510-2; Sequence=VSP_007106, VSP_007107, VSP_007108;
CC -!- TISSUE SPECIFICITY: Very sparingly expressed, although is detected
CC at low levels in testis, prostate, ovary, heart, placenta, lung
CC and pancreas. Overexpressed in several brain, colon and breast
CC carcinomas.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF488804; AAC15766.1; -.
CC EMBL: AJ515153; CAD56159.3; -.
CC EMBL: AJ515154; CAD56160.2; -.
CC Genew; HGNC:117178; ADAMTS20.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept M Zn BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSPI.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.
CC SMART; SM00209; TSPI; 12.
CC PROSITE; PS00215; ADAM_MEPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00092; TSPI; 12.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix; Alternative splicing.
CC SIGNAL 1 21
CC PROPEP 22 253
CC CHAIN 254 1911
CC DOMAIN 254 467
CC DOMAIN 468 555
CC DOMAIN 556 611
CC DOMAIN 612 723
CC DOMAIN 724 846
CC DOMAIN 847 905
CC DOMAIN 906 962
CC DOMAIN 967 1024
CC DOMAIN 1025 1074
CC DOMAIN 1077 1136
CC DOMAIN 1153 1207
CC DOMAIN 1208 1265
CC DOMAIN 1305 1357
CC DOMAIN 1359 1417
CC DOMAIN 1418 1476
CC DOMAIN 1477 1532
CC DOMAIN 1536 1589
CC DOMAIN 1590 1653
CC DOMAIN 1655 1711
CC DOMAIN 1712 1911
CC GON.
```

FT	METAL	403	403	ZINC (CATALYTIC) (BY SIMILARITY).	DE	metalloproteinase with thrombospondin motifs 20 (ADAM-TS 20) (ADAM-TS20).
FT	ACT SITE	404	404	BY SIMILARITY.	DE	ADAMTS20.
FT	METAL	407	407	ZINC (CATALYTIC) (BY SIMILARITY).	OS	Mus musculus (Mouse).
FT	METAL	413	413	ZINC (CATALYTIC) (BY SIMILARITY).	OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
FT	CARBOHYD	92	92	N-LINKED (GLCNAC. . .) (POTENTIAL).	OC	Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).	OX	NCBI_TaxID=10090;
FT	CARBOHYD	445	445	N-LINKED (GLCNAC. . .) (POTENTIAL).	RN	[1]
FT	CARBOHYD	702	702	N-LINKED (GLCNAC. . .) (POTENTIAL).	RP	SEQUENCE FROM N.A. (ISOFORM 1).
FT	CARBOHYD	717	717	N-LINKED (GLCNAC. . .) (POTENTIAL).	RC	TISSUE=Petal brain;
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. . .) (POTENTIAL).	RX	MEDLINE=22566039; PubMed=12562771;
FT	CARBOHYD	809	809	N-LINKED (GLCNAC. . .) (POTENTIAL).	RA	Llamazares M., Cal S., Quesada V., Lopez-Otin C.;
FT	CARBOHYD	870	870	N-LINKED (GLCNAC. . .) (POTENTIAL).	RT	"Identification and characterization of ADAMTS-20 defines a novel
FT	CARBOHYD	1062	1062	N-LINKED (GLCNAC. . .) (POTENTIAL).	RT	subfamily of metalloproteinases-disintegrins with multiple
FT	CARBOHYD	1457	1457	N-LINKED (GLCNAC. . .) (POTENTIAL).	RL	thrombospondin-1 repeats and a unique GON domain.";
FT	CARBOHYD	1543	1543	N-LINKED (GLCNAC. . .) (POTENTIAL).	RL	J. Biol. Chem. 278:13382-13389 (2003).
FT	CARBOHYD	1573	1573	N-LINKED (GLCNAC. . .) (POTENTIAL).	RN	[2]
FT	CARBOHYD	1764	1764	N-LINKED (GLCNAC. . .) (POTENTIAL).	RP	SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.
FT	CARBOHYD	1782	1782	N-LINKED (GLCNAC. . .) (POTENTIAL).	RC	STRAIN=DBA/2;
FT	CARBOHYD	1853	1853	N-LINKED (GLCNAC. . .) (POTENTIAL).	RX	MEDLINE=22806432; PubMed=12925592;
FT	VARSPLIC	683	693	THD1CVOGQCM -> SYNIDCNVLIK (in isoform 2).	RA	Rao C., Foernzler D., Loftus S.K., Liu S., McPherson J.D.,
FT				/FtId=VSP_007106.	RA	Jungers K.A., Apte S.S., Pavan W.J., Beier D.R.;
FT	VARSPLIC	1430	1504	CSACGGRKRYEVICDQFQKLEDTNCSQVQKPTKAC	RT	"A defect in a novel ADAMTS family member is the cause of the belted
FT				RSVRCPSWKANSWNECVTCGSGVQOQVYCRLK -> EDL	RT	white-spotting mutation.";
FT				KVKLLPQRTIILWELMKNIFCHGKSHVLYLVNVDHLLYP	RL	Development 130:4665-4672 (2003).
FT				RHCPETETIETFLSLSQFTWGLDKYKNSL (in isoform 2).	CC	-!- FUNCTION: May play a role in tissue-remodeling process occurring
FT				/FtId=VSP_007107.	CC	in both normal and pathological conditions.
FT				Missing (in isoform 2).	CC	-!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
FT				/FtId=VSP_007108.	CC	-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
FT				T -> V (IN REF. 2).	CC	matrix (By similarity).
FT	CONFLICT	152	152	E -> Y (IN REF. 2).	CC	-!- ALTERNATIVE PRODUCTS:
FT	CONFLICT	360	360	T -> M (IN REF. 2).	CC	Event=Alternative splicing; Named isoforms=2;
FT	CONFLICT	371	371	D -> E (IN REF. 2).	CC	Name=1; Synonyms=ADAMTS20 B long isoform;
FT	CONFLICT	456	456	I -> V (IN REF. 2).	CC	Isoid=P59511-1; Sequence=displayed;
FT	CONFLICT	648	648	MISSING (IN REF. 2).	CC	Name=2; Synonyms=ADAMTS20 A short isoform;
FT	CONFLICT	792	793	LILQ -> ILIE (IN REF. 2).	CC	Isoid=P59511-2; Sequence=VSP_007606; VSP_007607;
FT	CONFLICT	818	821	GMLLAK -> WHVIG (IN REF. 2).	CC	-!- TISSUE SPECIFICITY: Expressed at low level in testis and brain.
FT	CONFLICT	908	913	S -> Q (IN REF. 2).	CC	-!- PTM: The precursor is cleaved by a furin endopeptidase (By
FT	CONFLICT	1316	1316	T -> R (IN REF. 2).	CC	similarity).
FT	CONFLICT	1328	1328	T -> A (IN REF. 2).	CC	-!- DISEASE: Defects in ADAMTS20 are the cause of the belted (bt)
FT	CONFLICT	1817	1817	E -> Q (IN REF. 2).	CC	phenotype. It is a pigmental defect which occurs as a result of a
FT	CONFLICT	1882	1882	CF592E220D32B250 CRC64;	CC	defect in melanocyte development.
FT	SEQUENCE	1911	AA; 214656 MW; CF592E220D32B250 CRC64;		CC	-!- SIMILARITY: Belongs to peptidase family M12B.
QY				Query Match	CC	-!- SIMILARITY: Contains 1 disintegrin-like domain.
Db				Best Local Similarity	CC	-!- SIMILARITY: Contains 1 GON domain.
QY				Matches 71; Conservative 36; Mismatches 97; Indels 6; Gaps 2;	CC	-!- SIMILARITY: Contains 15 TSP type-1 domains.
Db				7 HLELLVAGDPVFOAHQEDTTRYVLTNIGAELEDFSLGQFRVHLVQWVILTEPEGA 66	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
Db				260 YIEIMVTADAKVSAHGSNLQNYIILMSIVATIKDPISGNLIHIVVVKLVMIHREERG 319	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
QY				67 PNITANLTSSLLSVCGWSQTNPEDDTPGHADLVLYITRPDLPLDGNRQVRGVTLGG 126	CC	the European Bioinformatics Institute. There are no restrictions on its
Db				320 FVINFGATTAKNFCSWQQTQNDLVHPSHHDTAVLTREDICSKKCNLTGSLYLT 379	CC	use by non-profit institutions as long as its content is in no way
QY				127 ACSPWTSCLITGDFGLGVITIAEIGHGSFGLHGDGPGSGGG---PSCHVWASDGAAP 182	CC	modified and this statement is not removed. Usage by and for commercial
Db				380 ICDPLQSCFINEEKLISAFIAHLEGLTGLVQCHDNP--RCKEMKVTYKHVWAPALSH 437	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
QY				183 RAGLWSPFCRRQLLSLSAGRCARCVWDPP 212	CC	or send an email to license@isb-sib.ch).
Db				438 MSPSWNSWNSCRKYTFEFLDTGVECLLDKP 467	CC	-----
RESULT 4					DR	EMBL; AJ512753; CAD54808.3; -
ID_20_MOUSE					DR	EMBL; AY189815; AA074895.1; -
AC					DR	EMBL; AY189816; AA074896.1; -
DT					DR	MGI; MGI:2660628; Adamts20.
DT					DR	InterPro; IPR001762; Disintegrin.
DT					DR	InterPro; IPR006025; Pept_M_Zn_BS.
DT					DR	InterPro; IPR002870; Peptidase_M12B_N.
DT					DR	InterPro; IPR000884; TSP1.
DE					DR	Pfam; PF01562; Pep_M12B_propep; 1.
					DR	Pfam; PF01421; Repolysin; 1.
					DR	Pfam; PF00090; tsp_1; 11.
					DR	SMART; SM00209; TSP1; 14.
					DR	PROSITE; PS00215; ADAM_MEPRO; 1.
					DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
					DR	PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
					DR	PROSITE; PS00092; TSP1; 13.
					DR	PROSITE; PS00142; ZINC_PROTEASE; 1.

DR	PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.	OX	NCBI_TaxID=9606;
DR	PROSITE; PS50900; PLAC; 1.	RN	[1]
DR	PROSITE; PS50092; TSP1; 5.	RP	SEQUENCE FROM N.A.
DR	PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.	RA	Casas C., Pritchard M.A., Estivill X., Arbones M.L.;
KW	Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;	RT	"Cloning, characterization and mapping on human chromosome 21 of the
KW	Repeat; Extracellular matrix; Alternative splicing.	RL	orthologue of murine Adams-1.";
FT	SIGNAL 1 24	RL	Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
FT	POTENTIAL	[2]	
FT	PROPEP 25 279	RN	SEQUENCE FROM N.A., AND FUNCTION.
FT	CHAIN 280 1224	RP	TISSUE=Heart;
FT	DOMAIN 280 495	RC	MEDLINE=99367466; PubMed=10438512;
FT	DOMAIN 496 585	RX	Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
FT	DOMAIN 586 641	RA	Lombardo M., Iruela-Arispe M.L.;
FT	DOMAIN 642 746	RA	"METH-1, a human ortholog of ADAMS-1, and METH-2 are members of a new
FT	DOMAIN 747 873	RT	family of proteins with angio-inhibitory activity.";
FT	DOMAIN 874 922	RL	J. Biol. Chem. 274:23349-23357(1999).
FT	DOMAIN 923 987	RN	[3]
FT	DOMAIN 988 1048	RP	SEQUENCE FROM N.A.
FT	DOMAIN 1051 1115	RC	TISSUE=Endothelial cells;
FT	DOMAIN 1127 1181	RX	MEDLINE=20247184; PubMed=10785405;
FT	DOMAIN 1186 1223	RA	Glenke J., Schmitt A.O., Pilarsky C., Hinzmann B., Weiss B.,
FT	SITE 249 249	RA	Rosenthal A., Thierauch K.H.;
FT	METAL 433 433	RT	"Differential gene expression by endothelial cells in distinct
FT	ACT_SITE 434 434	RL	angiogenic states.";
FT	METAL 437 437	RL	Eur. J. Biochem. 267:2820-2830(2000).
FT	METAL 443 443	[4]	
FT	CARBOHYD 156 156	RP	SEQUENCE FROM N.A.
FT	CARBOHYD 310 310	RC	TISSUE=Brain;
FT	CARBOHYD 741 741	RX	MEDLINE=20181126; PubMed=10718198;
FT	CARBOHYD 780 780	RA	Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
FT	CARBOHYD 835 835	RT	"Prediction of the coding sequences of unidentified human genes. XVI.
FT	CARBOHYD 905 905	RT	The complete sequences of 150 new cDNA clones from brain which code
FT	CARBOHYD 935 935	RT	for large proteins in vitro.";
FT	VARSPLIC 1063 1072	RL	DNA Res. 7:65-73(2000).
FT	VARSPLIC 1073 1224	[5]	
FT	CONFLICT 284 284	RP	SEQUENCE FROM N.A.
FT	SEQUENCE 1224 AA; 136186 MW; 572E72129E1401F8 CRC64;	RX	MEDLINE=20289799; PubMed=10830953;
Query Match	27.0%; Score 307.5; DB 1; Length 1224;	RA	Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
Best Local Similarity	36.2%; Pred. No. 1.8e-18;	RA	Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
Matches 77; Conservative 35; Mismatches 92; Indels 9; Gaps 5;		RA	Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
		RA	Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
		RA	Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
		RA	Rosenthal A., Kudon J., Shibuya K., Kawasaki K., Asakawa S.,
		RA	Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
		RA	Minoshima S., Shiohara N., Nordstiek G., Hornischer K., Brandt P.,
		RA	Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
		RA	Ramer J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
		RA	Wenmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
		RA	Lehrach H., Reinhardt K., Yaspo M.-L.;
		RT	"The DNA sequence of human chromosome 21.";
		RL	Nature 405:311-319(2000).
		[6]	
		RN	SEQUENCE OF 418-967 FROM N.A.
		RP	TISSUE=Melanoma;
		RC	Blum H., Bauersachs S., Mewes H.-W., Weil B., Wienann S.;
		RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
		CC	-I- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
		CC	involved in its turnover (By similarity). Has angiogenic inhibitor
		CC	activity. Active metalloprotease, which may be associated with
		CC	various inflammatory processes as well as development of cancer
		CC	cachexia. May play a critical role in follicular rupture.
		CC	-I- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1938-Glu-1939
		CC	site, within the chondroitin sulfate attachment domain.
		CC	-I- COPACITOR: Binds 1 zinc ion per subunit (By similarity).
		CC	-I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
		CC	matrix (By similarity).
		CC	-I- DOMAIN: The spacer domain and the TSP type-1 domains are important
		CC	for a tight interaction with the extracellular matrix.
		CC	-I- PTM: The precursor is cleaved by a furin endopeptidase (By
		CC	similarity).
		CC	-I- SIMILARITY: Belongs to peptidase family M12B.
		CC	-I- SIMILARITY: Contains 1 disintegrin-like domain.
		CC	-I- SIMILARITY: Contains 3 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF170084; AAF15317.1; -
CC EMBL; AF060152; AAD48080.1; ALT_INIT.
CC EMBL; AF207664; AAF23772.1; -
CC EMBL; AB037767; BAA92584.1; ALT_INIT.
CC EMBL; AP001697; BAA95502.1; -
CC EMBL; ALI62080; CAB82413.1; -
CC PIR; T47158; T47158.
CC MEROPS; M12.222; -
CC Genew; HGNC:217; ADAMTS1.
CC MIM; 605174; -
CC GO; GO:0008237; Fmetallopeptidase activity; TAS.
CC GO; GO:0007229; P; integrin-mediated signaling pathway; TAS.
CC GO; GO:0008285; P; negative regulation of cell proliferation; TAS.
CC InterPro; IPR006586; ADAM cysteine.
CC InterPro; IPR001782; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 3.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00608; ACE; 1.
CC SMART; SM00209; TSP1; 3.
CC PROSITE; PS00215; ADAM_MEPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 49 POTENTIAL.
FT PROPEP 50 252 BY SIMILARITY.
FT CHAIN 253 967 ADAMTS1.
FT DOMAIN 253 475 METALLOPROTEASE.
FT DOMAIN 476 559 DISINTEGRIN-LIKE.
FT DOMAIN 559 614 TSP TYPE-1 1.
FT DOMAIN 617 724 CYS-RICH.
FT DOMAIN 725 849 SPACER.
FT DOMAIN 854 905 TSP TYPE-1 2.
FT DOMAIN 908 967 TSP TYPE-1 3.
FT DOMAIN 843 846 POLY-LYS.
FT SITE 198 198 CYS-STEIN SWITCH (POTENTIAL).
FT METAL 401 401 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 402 402 BY SIMILARITY.
FT METAL 405 405 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 411 411 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 227 227 P -> A (IN REF. 4 AND 5).
FT CONFLICT 468 468 P -> H (IN REF. 1).
FT CONFLICT 561 561 S -> N (IN REF. 1).
SQ SEQUENCE 967 AA; C189389324741ED1 CRC64;
Query Match 26.7%; Score 304.5; DB 1; Length 967;
Best Local Similarity 29.8%; Pred. No. 2.4e-18;
Matches 65; Conservative 40; Mismatches 94; Indels 19; Gaps 5;
7 HLELLAVGPDVQAHQEDTERVLTNLNIGAEILLRDPSELGQAFRVHLVQVILTEPEGA 66
259 YVETMLVDQSMAEFFHGSLGKHLTLTLFVAARLYKHPISIRNSVSLVVKILVHDEQKG 318

QY 67 PNITANTLSLLSCVGSQTINBEDDDTPGHADLVLYITRDELELPDGNR--QVRGVTLQ 124
DB 319 PEVTSNAALTRNFCNQKQHNPPSDRDAHYDTAILFTQDL---CGSQTCDTIGMADV 375
QY 125 GGACSPWTSCLITGDFGLGVITIAEIGHISFGSEHGDGAPG----SGCGSPSGHVMAAS--- 177
DB 376 GTVCDPSRSCSVIEDDGLQAFTHAELGHVFNPHDDAKQACASLNGVQDSHVMASMLS 435
QY 178 --DGAAPRAGLAWSPCRRROLLSLSRAGRVCVWDPPR 213
DB 436 NLDHSQP----WSPCSAYMITSLFDNGHGCELDKMPQ 468
RESULT 7
AT10 HUMAN
ID AT10 HUMAN STANDARD; PRT; 1077 AA.
AC Q9H324;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-10 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 10) (ADAM-TS 10) (ADAM-TS10) (Fragment).
GN ADAMTS10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Apte S.S.;
RT "ADAM-TS10: a novel member of the ADAM-TS family containing multiple
RT thrombospondin type I repeats."; to the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AF163762; AAG35563.1; -
CC MEROPS; M12.235; -
CC Genew; HGNC:13201; ADAMTS10.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 5.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 5.
CC PROSITE; PS00215; ADAM_MEPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00090; TSP1; 5.
CC PROSITE; PS00092; TSP1; 5.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.

```
KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT NON_TER 1
FT PROPEP 1 207
FT CHAIN 208 1077
FT DOMAIN 209 433
FT DOMAIN 434 520
FT DOMAIN 578 679
FT DOMAIN 680 802
FT DOMAIN 521 576
FT DOMAIN 799 857
FT DOMAIN 858 919
FT DOMAIN 921 975
FT DOMAIN 977 1032
FT DOMAIN 1039 1077
FT METAL 366 367
FT ACT_SITE 367 367
FT METAL 370 370
FT METAL 376 376
FT CARBOHYD 64 64
FT CARBOHYD 196 196
FT CARBOHYD 297 297
FT CARBOHYD 714 714
FT CARBOHYD 769 769
FT CARBOHYD 866 866
SQ SEQUENCE 1077 AA; 118072 MW; 3914DE18DCBBF587 CRC64;

Query Match 26.58; Score 302.5; DB 1; Length 1077;
Best Local Similarity 33.5%; Pred. No. 4.1e-18;
Matches 75; Conservative 37; Mismatches 91; Indels 21; Gaps 7;

QY 7 HLELLVAVGPDVQAH-QEDTERVVLNLTNGAELLRDPESLGAQPRVHLKXNVILTEPEG 65
DB 214 YVETLVADKQMVAYHQRDVEQVLAIMNIVAKLFDQSSLSGSTVNIIVRLILITEDQP 273

QY 66 APNITANLTSSLSVCGWSQTI-----NPEDDTPGADLVHYITRFDL-----ELPDG 114
DB 274 TLEITHAGKSLDSFCWKQKSVNHSCHGNAIPENGVAHDTAVLITRYDICIYKPKCG 333

QY 115 NROVRGVTQLGACSPWCSLITDTEGDLGVTIAETGHSFGLBHDGAPSGCGPSGH- 173
DB 334 ---TLGLAPVGGCGERERSCSVNEIDIGLPQAFTHAIEIGHTFGMNHGTV-GNSCGARGOD 389

QY 174 ---VMSADGAAPRAGLAWSPCRQLLSLLSAGARCVCWD-PPR 213
DB 390 PAKLMAAHITKTNPFVWSSCNRDYITSFDSGLGLCLNRRPPR 433

RESULT 8
ATSI_RAT STANDARD; PRT; 967 AA.
AC Q9WUQ1; Q9ER11;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Brain;
RA Liu X., Fu Y., Yin T., Johnstone E.M., Stephenson D.T., Clemens J.A.,
RA Little S.P.;
RT "Induction of a disintegrin and metalloprotease with the
RT thrombospondin type I motif (ADAMTS).";
RL Submitted (MAY-1999) to the ENBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 18-967 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;

RX MEDLINE=20304099; PubMed=10847486;
RA Diamantis I., Luethi M., Hoesli M., Reichen J.;
RT "Cloning of the rat ADAMTS-1 gene and its down regulation in
RT endothelial cells in cirrhotic rats.";
RL Liver 20:165-172(2000).
CC !- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. Has angiogenic inhibitor activity (By
CC similarity). Active metalloprotease, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture (By
CC similarity).
CC !- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1683-Glu-|-Leu-1684
CC site, within the chondroitin sulfate attachment domain.
CC !- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC !- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC !- INDUCTION: Down-regulated in endothelial cells derived from
CC cirrhotic liver.
CC !- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC !- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC !- SIMILARITY: Belongs to peptidase family M12B.
CC !- SIMILARITY: Contains 1 disintegrin-like domain.
CC !- SIMILARITY: Contains 3 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; AF149118; AAC34012.1; -
CC EMBL; AF304446; AAC29823.1; -
CC MEROPS; M12.222; -
CC InterPro; IPR006586; ADAM cysteine.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC Pfam; PF01562; Pep_M12B_Proprep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp 1; 3.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SMC0608; ACR; 1.
CC SMART; SMC0209; TSP1; 3.
CC PROSITE; PS0215; ADAM_MERPO; 1. FALSE NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC PROSITE; PS00092; TSP1; 3.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 54
FT SIGNAL 1 54
FT PROPEP 55 252
FT CHAIN 253 967
FT DOMAIN 253 475
FT DOMAIN 476 558
FT DOMAIN 559 614
FT DOMAIN 616 724
FT DOMAIN 725 857
FT DOMAIN 854 910
FT DOMAIN 911 967
FT DOMAIN 194 198
FT SITE 205 205
FT METAL 401 401
FT ACT_SITE 402 402
FT METAL 405 405
FT METAL 411 411
```

FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 764 764 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 945 945 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 21 21 I -> V (IN REF. 2).
FT CONFLICT 26 31 KFRSSQ -> RSRGSL (IN REF. 2).
FT CONFLICT 49 49 V -> A (IN REF. 2).
FT CONFLICT 72 72 R -> P (IN REF. 2).
FT CONFLICT 79 79 L -> TR (IN REF. 2).
FT CONFLICT 249 249 R -> G (IN REF. 2).
FT CONFLICT 262 265 TMIV -> NLK (IN REF. 2).
FT CONFLICT 607 607 S -> F (IN REF. 2).
FT CONFLICT 936 936 L -> V (IN REF. 2).
FT CONFLICT 962 962 I -> T (IN REF. 2).
SQ SEQUENCE 967 AA; 105705 MW; F93C86F6DCDB4CF CRC64;
Query Match 26.2%; Score 299.5; DB 1; Length 967;
Best Local Similarity 29.6%; Pred. No. 6.5e-18;
Matches 64; Conservative 36; Mismatches 101; Indels 15; Gaps 4;
QY 7 HLELLVAVGPDVFAHQEDERYVLTNLNTGIAELLRPDSLGAQFRVHLVQVILTEPEGA 66
Db 259 YVETMLVADQSMADFHGSLKHVLLTTFVSAARFYKPSIRNSISLVVVKLVIVVEQKG 318
QY 67 PNITANLTSLLSVCGMSQTNPDPTDGHADLVLYITRFDLELDPGNRQVRGVTLGG 126
Db 319 PEVTSNAALTNRNFCSSQKQNSPSDRDEHYDTAILFTRODL-CGSHTCDTLGNADVGT 377
QY 127 ACPTWSCSLTEBTGPDGLVGTIAHEIGHSFGLHDGAPG-----SCGSPGHVMS----- 177
Db 378 VCDPSRSCSVIEDGLQAQPTTAHELGHVFNPHDDAKHCAFNGSVGSDSHLMASLSSL 437
QY 178 DGAAPRAGLAWSPCSRQLLSLSAGRARCVWDPPR 213
Db 438 DRSQP-----WSPCSAYMVSFLDNGHGCECLMDKPO 468
RESULT 9
ATSI1_MOUSE
ID ATSI1_MOUSE STANDARD; PRT; 968 AA.
AC F97857; OS4768;
DT 30-MAY-2000 (Rel. 39, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-1 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 1) (ADAM-TS 1) (ADAM-TS1).
GN ADAMTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=98110583; PubMed=9441751;
RA Kuno K., Litasa H., Ohno S., Matsushima K.;
RT "The exon/intron organization and chromosomal mapping of the mouse
RT ADAMTS-1 gene encoding an ADAM family protein with TSP motifs.";
RL Genomics 46:466-471(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97150761; PubMed=9995297;
RA Kuno K., Kanada N., Nakashima E., Fujiki F., Ichimura F.,
RA Matsushima K.;
RT "Molecular cloning of a gene encoding a new type of metalloproteinase-
RT disintegrin family protein with thrombospondin motifs as an
RT inflammation associated gene";
RL J. Biol. Chem. 272:556-562(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Limb, and Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish M.A.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP CHARACTERIZATION, AND MUTAGENESIS OF GLU-403.
RX MEDLINE=99303657; PubMed=10373500;
RA Kuno K., Terashima Y., Matsushima K.;
RT "ADAMTS-1 is an active metalloproteinase associated with the
RT extracellular matrix.";
RL J. Biol. Chem. 274:18821-18826(1999).
RN [5]
RP FUNCTION.
RX MEDLINE=20389568; PubMed=10930576;
RA Kuno K., Okada Y., Kawashima H., Nakamura H., Miyasaka M.,
RA Ohno H., Matsushima K.;
RT "ADAMTS-1 cleaves a cartilage proteoglycan, aggrecan.";
RL FEBS Lett. 478:241-245(2000).
RN [6]
RP FUNCTION, AND INDUCTION
RX MEDLINE=20243757; PubMed=10781075;
RA Robker R.L., Russell D.L., Espey L.L., Lydon J.P., O'Malley B.W.,
RA Richards J.S.;
RT "Progesterone-regulated genes in the ovulation process: ADAMTS-1 and
RT cathepsin L proteases";
RL Proc. Natl. Acad. Sci. U.S.A. 97:4689-4694(2000).
CC -I- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. Has antigenic inhibitor activity (By
CC similarity). Active metalloproteinase, which may be associated with
CC various inflammatory processes as well as development of cancer
CC cachexia. May play a critical role in follicular rupture (By
CC similarity).
CC -I- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1691-Glu-Leu-1692
CC site, within the chondroitin sulfate attachment domain.
CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix.
CC -I- INDUCTION: Induced in vitro in colon adenocarcinoma cells by
CC interleukin-1, or in vivo in kidney and heart by
CC lipopolysaccharide. Also induced by LH stimulation in granulosa
CC cells of preovulatory follicles.
CC -I- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -I- PTM: The precursor is cleaved by a furin endopeptidase.
CC -I- SIMILARITY: Belongs to peptidase family M128.
CC -I- SIMILARITY: Contains 1 disintegrin-like domain.
CC -I- SIMILARITY: Contains 3 TSP type-1 domains.
CC -I- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF261918; AAF89106.1; -
 CC EMBL; AF488803; AAO15765.1; -
 CC EMBL; AB037733; BAA92550.1; -
 CC HSSP; P15167; IATL.
 CC MEROPS; M12.021; -
 CC Genew; HGNC:13202; ADAMTS9.
 CC MIM; 603421; -
 CC GO; GO:0008237; F:metalloproteinase activity; TAS.
 CC GO; GO:0007275; P:development; TAS.
 CC GO; GO:0006516; P:glycoprotein catabolism; TAS.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR001818; Pept_M12B_M12B.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR001590; Peptidase_M12B.
 CC InterPro; IPR002870; Peptidase_M12B_N.
 CC InterPro; IPR000884; TSP1.
 CC Pfam; Pf01562; Pfam_M12B_propep; 1.
 CC Pfam; Pf01421; Repolysin; 1.
 CC Pfam; PF00090; tsp1; 11.
 CC SMART; SM00209; TSP1; 12.
 CC PROSITE; PS00215; ADAM_MEROP; 1.
 CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 CC PROSITE; PS00092; TSP1; 14.
 CC PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Alternative splicing.
 FT SIGNAL 1 18
 FT PROPEP 19 287
 FT CHAIN 288 1935
 FT DOMAIN 293 499
 FT DOMAIN 509 587
 FT DOMAIN 588 643
 FT DOMAIN 644 752
 FT DOMAIN 753 877
 FT DOMAIN 878 936
 FT DOMAIN 939 997
 FT DOMAIN 998 1049
 FT DOMAIN 1052 1109
 FT DOMAIN 1110 1166
 FT DOMAIN 1182 1240
 FT DOMAIN 1241 1296
 FT DOMAIN 1328 1379
 FT DOMAIN 1382 1440
 FT DOMAIN 1441 1494
 FT DOMAIN 1497 1555
 FT DOMAIN 1556 1611
 FT DOMAIN 1612 1676
 FT DOMAIN 1677 1734
 FT DOMAIN 1735 1935
 FT SITE 88 96
 FT SITE 223 223
 FT METAL 434 434
 FT ACT_SITE 435 435
 FT METAL 438 438
 FT METAL 444 444
 FT CARBOHYD 112 112
 FT CARBOHYD 135 135
 FT CARBOHYD 271 271
 FT CARBOHYD 749 749
 FT CARBOHYD 840 840
 FT CARBOHYD 1213 1213
 FT CARBOHYD 1267 1267

FT CARBOHYD 1788 1788 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 1806 1806 N-LINKED (GLCNAC...) (POTENTIAL).
 FT VARSPLIC 1064 1072 CLVTGKGH->VRWEGCYFP (in isoform 3).
 FT VARSPLIC 1073 1935 /FTid=VSP_005499.
 FT VARSPLIC 1624 1629 Missing (in isoform 3).
 FT VARSPLIC 1630 1935 /FTid=VSP_005500.
 FT VARSPLIC 1630 1935 CSVTGCG->VPSWEL (in isoform 2).
 FT VARSPLIC 1630 1935 /FTid=VSP_007548.
 FT VARSPLIC 1630 1935 Missing (in isoform 2).
 FT VARSPLIC 1630 1935 /FTid=VSP_007549.
 FT CONFLICT 46 46 S->G (IN REF. 1).
 FT CONFLICT 96 96 P->S (IN REF. 1).
 FT CONFLICT 182 182 D->G (IN REF. 2).
 FT CONFLICT 367 367 P->L (IN REF. 1).
 FT CONFLICT 1117 1117 V->G (IN REF. 3).
 SQ SEQUENCE 1935 AA; 216556 MW; FD3D51B88300A3C6 CRC64;
 Query Match 25.6%; Score 292; DB 1; Length 1935;
 Best Local Similarity 31.0%; Pred. No. 6.4e-17;
 Matches 67; Conservative 38; Mismatches 95; Indels 16; Gaps 5;
 QY 8 LELLVAVGPDVFOAHQEDTERVLTNINIGALLRDPGLGAGQFRVHLVXVILTEPEGAP 67
 Db 295 VEVLVADNMVSYHGENLQHYIILMSIVASYKDPISIGNLINIVNLVHNEQDGP 354
 QY 68 NITANLTSSLVCGWSQTINPEDDTPG--HADLVLYITPDLPLDGNRQVRGVTLG 125
 Db 355 SISFNAQTLLKNFCQWQHSKN-----SPGGIHHDTAVLLTRDICIHAHDKCDTGLAELG 409
 QY 126 GACSPWMSCLITDTPGLGVTHAIEHCHSFGLEHDPAGSGCGPSG-----HVMASDGA 180
 Db 410 TICDPYRSCISIDSGISTAFTHAIEHGHVFNMPHD--DNNKKEEGVKSPQHVMAPTLN 467
 QY 181 APRAGLAWSPCSRQLLSAGRCARCVWDPP--RP 214
 Db 468 FYTNPMWMSKSKYITEFLDTGYGECULNPESEPP 503
 RESULT 11
 AT88 HUMAN STANDARD; PRT; 890 AA.
 AC Q9UP79; Q9NZS0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with chromospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)
 DE (METH-8).
 GN ADAMTS8 OR METH2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=99367466; PubMed=10438512;
 RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
 RA Lombardo M., Iruela-Arispe M.L.;
 RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
 RT family of proteins with angio-inhibitory activity.";
 RL J. Biol. Chem. 274:23349-23357(1999).
 RN [2]
 RP SEQUENCE OF 195-440 FROM N.A.
 RX MEDLINE=20079168; PubMed=10610729;
 RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
 RA "ADAM-TS8, a novel metalloproteinase of the ADAM-TS family located on
 RT mouse chromosome 9 and human chromosome 11.";
 RL Genomics 62:312-315(1999).
 CC -/- FUNCTION: Has anti-angiogenic properties.
 CC -/- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -/- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).

CC	-!	TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower expression in brain, placenta, heart, stomach and fetal brain and kidney.
CC	-!	DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.
CC	-!	PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
CC	-!	SIMILARITY: Belongs to peptidase family M12B.
CC	-!	SIMILARITY: Contains 1 disintegrin-like domain.
CC	-!	SIMILARITY: Contains 2 TSP type-1 domains.
CC	-!	This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC	EMBL	AF060153; AAD48081.1; -.
CC	EMBL	AF175283; AAF25806.1; -.
DR	HSSP	P34179; 1IAG.
DR	MEROPS	M12.226; -.
DR	Genew	HGNC:224; ADAMTS8.
DR	MIM	605175; -.
DR	GO	GO:0005178; F.integrin binding; TAS.
DR	GO	GO:0008237; F.metalloproteinase activity; TAS.
DR	GO	GO:0007345; P.negative regulation of cell proliferation; TAS.
DR	GO	GO:0008285; P.negative regulation of cell proliferation; TAS.
DR	InterPro	IPR001762; Disintegrin.
DR	InterPro	IPR006025; Pept_Mn_Es.
DR	InterPro	IPR001590; Peptidase_M12B.
DR	InterPro	IPR002870; Peptidase_M12B_N.
DR	InterPro	IPR000884; TSP1.
DR	InterPro	IPR008085; TSP 1.
DR	Pfam	PF01562; Pept_M12B_prop; 1.
DR	Pfam	PF01421; Reprolysin; 1.
DR	Pfam	PF00090; tsp 1; 2.
DR	PRINTS	PRO1705; TSP1REPEAT.
DR	SMART	SM00209; TSP1 2.
DR	PROSITE	PS50215; ADAM_MEROP; 1.
DR	PROSITE	PS00427; DISINTEGRIN_1; FALSE_NEG.
DR	PROSITE	PS50214; DISINTEGRIN_2; FALSE_NEG.
DR	PROSITE	PS50092; TSP1 2.
DR	PROSITE	PS00142; ZINC_PROTEASE; 1.
DR	Hydrolase	Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW	Repeat	Extracellular matrix; Heparin-binding.
FT	SIGNAL	1 27 POTENTIAL.
FT	PROPEP	28 214 BY SIMILARITY.
FT	CHAIN	215 890 ADAMTS-8.
FT	DOMAIN	215 430 METALLOPROTEASE.
FT	DOMAIN	439 526 DISINTEGRIN-LIKE.
FT	DOMAIN	527 582 TSP TYPE-1 1.
FT	DOMAIN	584 690 CYS-RICH.
FT	DOMAIN	691 832 TSP TYPE-1 2.
FT	DOMAIN	834 889 POLY-PRO.
FT	DOMAIN	202 205 ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	364 364 BY SIMILARITY.
FT	ACT_SITE	365 365 ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	368 368 ZINC (CATALYTIC) (BY SIMILARITY).
FT	METAL	374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	491 491 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	600 600 E -> R (IN REF. 2).
FT	CONFLICT	195 195 YLTLLDGGHGDCLLDAPGAALPLPTGL -> FSGCHLQGM
FT	CONFLICT	413 440 IHFYLCVCVELKDLMP (IN REF. 2).
FT	SEQUENCE	890 AA; 96671 MW; 57D70BE03D5739D3 CRC64;
CC	Query Match	25.0%; Score 285; DB 1; Length 890;
CC	Best Local Similarity	29.9%; Pred. No. 1e-16;

Matches	63;	Conservative	36;	Mismatches	104;	Indels	8;	Gaps	3;
QY	8	LELVAVGPDVFAHQEDTERVYLTNLNIGAEILLRDPSPSLGAQFRVHLVKMVLTEPEGAP	67						
Db	222	VETLLVADASMAAFYAGDLQNHLILMSVAARIYKPSIKNSINLMVVKVLLIVDEKWP	281						
QY	68	NITANLTSSLLSVCGWSQTINPEDDTPDGHADLVLYITRFDLELPDGNRQVRGVTQLGGA	127						
Db	282	EVSNDGGLTLRNFQWQRFNQPSSDRHPEHYDPAILLTRQNFQCGEGLCDTLGVADIGTI	341						
QY	128	CSPTWSCIITEPTGDLGVITIAHEIOHSGLEHDPGSC-----GPSG--HVMASDGA	181						
Db	342	CDPNKSCSVIEDEGLQAHTLAHELGHVLSMPHDDSS--KPCTRLFGPMGKHVNAFLFVH	399						
QY	182	PRAGLANWPCSRRLQLLSLSSAGRCVWDPP	212						
Db	400	LNQTLFWSPCSAWLYTELJDGGHGDCLLDAP	430						
RESULT 12									
ATIS4 HUMAN		STANDARD;		PRT;		837 AA.			
AC	O75173; Q9UN83;								
DT	16-OCT-2001 (Rel. 40, Created)								
DT	16-OCT-2001 (Rel. 40, Last sequence update)								
DT	10-OCT-2003 (Rel. 42, Last annotation update)								
DE	ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1) (ADMP-1).								
GN	ADAMTS4 OR KIAA0688.								
OS	Homo sapiens (Human).								
OC	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
OX	NCBI_TaxID:9606;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Brain;								
RX	MEDLINE=98403880; PubMed=9734811;								
RA	Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,								
RA	Kotani H., Nomura N., Ohara O.;								
RT	"Prediction of the coding sequences of unidentified human genes. X. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro."								
RL	DNA Res. 5:169-176(1998).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=99286303; PubMed=10356395;								
RA	Tortorella M.D., Burn T.C., Pratta M.A., Abbaszade I., Hollis J.M.,								
RA	Liu R.-Q., Rosenfeld S.A., Copeland R.A., Decicco C.P., Wynn R.,								
RA	Rockwell A., Yang F., Duke J.L., Solomon K., George H., Bruckner R.,								
RA	Nagase H., Itoh Y., Ellis D.M., Ross H., Wiswall B.H., Murphy K.,								
RA	Hillman M.C. Jr., Hollis G.F., Newton R.C., Magolda R.L.,								
RA	Trzaskos J.M., Arner E.C.;								
RT	"Purification and cloning of aggrecanase-1: a member of the ADAMTS family of proteins."								
RL	Science 284:1664-1666(1999).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RA	Sawaji Y., Nagase H., Saklatvala J., Clark A.R.;								
RT	"ADAMTS-4 genomic locus."								
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBSJ databases.								
RN	[4]								
RP	PARTIAL SEQUENCE, AND CHARACTERIZATION.								
RX	MEDLINE=20400519; PubMed=10827174;								
RA	Tortorella M.D., Pratta M.A., Liu R.-Q., Abbaszade I., Ross H.,								
RA	Burn T.C., Arner E.C.;								
RT	"The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for aggrecan substrate recognition and cleavage."								
RL	J. Biol. Chem. 275:25791-25797(2000).								
CC	-!	FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be involved in its turnover. May play an important role in the destruction of aggrecan in arthritic diseases. Could also be a critical factor in the exacerbation of neurodegeneration in							

CC Alzheimer's disease. Cleaves aggrecan at the 392-Glu-|-Ala-393
CC site.
CC -|- CATALYTIC ACTIVITY: Glutamyl endopeptidase; bonds cleaved include
CC 370-Thr-Glu-Gly-Glu-|-Ala-Arg-Gly-Ser-377 in the interglobular
CC domain of mammalian aggrecan.
CC -|- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -|- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -|- TISSUE SPECIFICITY: Expressed in brain, lung and heart. Expressed
CC at very low level in placenta and skeletal muscles.
CC -|- INDUCTION: By interleukin-1.
CC -|- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -|- PTM: The precursor is cleaved by a furin endopeptidase.
CC -|- SIMILARITY: Belongs to peptidase family M12B.
CC -|- SIMILARITY: Contains 1 disintegrin-like domain.
CC -|- SIMILARITY: Contains 1 TSP type-1 domain.
CC -|- CAUTION: Has sometimes been referred to as ADAMTS2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB014588; BAA31663.1; -
CC EMBL; AF148213; RAD41494.1; -
CC EMBL; AY044847; AAL02262.1; -
CC PIR; T00355; T00355.
CC MEROPS; M12.221; -
CC Genew; HGNC:220; ADAMTS4.
CC MIM; 603876; -
CC GO; GO:0008237; F-metalloproteinase activity; TAS.
CC GO; GO:0006508; Proteolysis and peptidolysis; TAS.
CC GO; GO:0001501; P:skeletal development; TAS.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M12B.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR000884; TSP.
CC InterPro; IPR008085; TSP.
CC Pfam; PF01421; Repolysin; 1.
CC Pfam; PF00090; tsp_1; 1.
CC PRINTS; PR01705; TSP_REPEAT.
CC SMART; SM00209; TSP1; 1.
CC PROSITE; PS00215; ADAM_MPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00092; TSP1; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Extracellular matrix.
CC SIGNAL 1 51 POTENTIAL.
CC PROPEP 52 212
CC CHAIN 213 837
CC DOMAIN 213 436
CC DOMAIN 437 519
CC DOMAIN 520 575
CC DOMAIN 577 685
CC DOMAIN 685 837
CC DOMAIN 247 252
CC SITE 194 194
CC METAL 361 361
CC ACT_SITE 362 362
CC METAL 365 365
CC METAL 371 371
CC CARBOHYD 68 68
CC CONFLICT 77 77
CC CONFLICT 626 626
CC CONFLICT 682 682
CC SEQUENCE 837 AA; 5DF9C9AC137DF41F CRC64;

Query Match 24.9%; Score 284; DB 1; Length 837;
Best Local Similarity 32.5%; Pred. No. 1.2e-16;
Matches 69; Conservative 27; Mismatches 106; Indels 10; Gaps 4;
QY 8 LELLVAVGPDVFQAEQEDTERVYLNLTNGAELLRDPSPSLGAQPRVHLVKKWVILPEEGAP 67
Db 220 VETLVAVDDKMAAFEGAGLKYLLTWMAAAAKAFKPSIRNPVSLVTVRLVILSGEGRP 279
QY 68 NITANLTSLSSVCGWSQTINPEDDTPDGHADLVLTITFDLELPDGNRQVRGVYQLGGA 127
Db 280 QVGPSSAQTLSFCFAWQGLNTPEDSDPDHFTAILFTQDL-CGVSTCDTLGMDVGTV 338
QY 128 CSPWTSCLITDTPGLVGTIAEIGHSGLEHDGAPGSGC- - -GP- - -SGHYMASDGA 180
Db 339 CDPARSCAIVEDDGLQSAFTAHELGHVFNMLHDNS- -KPCISLNGPLSTRHWAPYMA 396
QY 181 APRAGLAWSPCSRRLQLLSLAGRAFCVWDPP 212
Db 397 HVDPEPNSPCSARFITDPLDNGYGHCLLDKP 428
RESULT 13
ID_ATS4 RAT STANDARD; PRT; 630 AA.
AC Q9ESP7; Q9ESP6; Q9ESP8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-4 precursor (EC 3.4.24.82) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 4) (ADAM-TS 4) (ADAM-TS4) (Aggrecanase 1)
DE (Fragment).
GN ADAMTS4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=20415831; PubMed=10961658;
RA Satch K., Suzuki N., Yokota H.;
RT ADAMTS-4 (a disintegrin and metalloproteinase with thrombospondin
RT motifs) is transcriptionally induced in beta-amyloid treated rat
RT astrocytes."
RL Neurosci. Lett. 289:177-180(2000).
CC -|- FUNCTION: Cleaves aggrecan, a cartilage proteoglycan, and may be
CC involved in its turnover. May play an important role in the
CC destruction of aggrecan in arthritic diseases. Cleaves aggrecan at
CC the 392-Glu-|-Ala-393 site.
CC -|- CATALYTIC ACTIVITY: Glutamyl endopeptidase; bonds cleaved include
CC 370-Thr-Glu-Gly-Glu-|-Ala-Arg-Gly-Ser-377 in the interglobular
CC domain of mammalian aggrecan.
CC -|- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -|- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -|- TISSUE SPECIFICITY: Only expressed in brain.
CC -|- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -|- PTM: The precursor is cleaved by a furin endopeptidase.
CC -|- SIMILARITY: Belongs to peptidase family M12B.
CC -|- SIMILARITY: Contains 1 disintegrin-like domain.
CC -|- SIMILARITY: Contains 1 TSP type-1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB042272; BAB16474.1; -
CC EMBL; AB042271; BAB16473.1; -

[illegible]

383 VTGCHPVRSTCLNHEDGSSAFVWAHETGHVLGMEHDG-QGNRCGDGDEVLGSIAPLVQ 441

181 APRAGLAWPCSRQQLLSLSSAGRACVMDPP 212

442 AAFERHWRSCSQQLSRYLHS--YDCLRDDP 471

RESULT 15

AT17 HUMAN STANDARD; PRT; 1095 AA.

AC Q8TE56; 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-17 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 17) (ADAM-TS 17) (ADAM-TS17).

GN ADAMTS17.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21856482; PubMed=1167212;

RA Cal S.; Obara A.J.; Llamazares M.; Garabaya C.; Quesada V.; Lopez-Otin C.;

FT "Cloning, expression analysis, and structural characterization of seven novel human ADAMTSs, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";

RT Gene 283:49-62(2002)

RL CC -! SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (by similarity).

CC -! TISSUE SPECIFICITY: Expressed in fetal lung, in adult brain, prostate, and liver.

CC -! PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -! SIMILARITY: Belongs to peptidase family M12B.

CC -! SIMILARITY: Contains 1 disintegrin-like domain.

CC -! SIMILARITY: Contains 1 PLAC domain.

CC -! SIMILARITY: Contains 5 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AJ315735; CAC86016.1; -.

DR MEROPS; M12.027; -.

DR Genew; HGNC:17109; ADAMTS17.

DR MIN; 607511; -.

DR InterPro; IPR001762; Disintegrin.

DR InterPro; IPR001818; Pept_M10A_M12B.

DR InterPro; IPR006025; Pept_M_2n_BS.

DR InterPro; IPR001590; Peptidase M12B.

DR InterPro; IPR002870; Peptidase_M12B_N.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.

DR Pfam; PF01562; Pep_M12B_propep; 1.

DR Pfam; PF01421; Reptolysin; 1.

DR Pfam; PF00090; tsp_1; 5.

DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00209; TSP1; 5.

DR PROSITE; PS0215; ADAM_MEPRO; 1.

DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.

DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.

DR PROSITE; PS0900; PLAC; 1.

DR PROSITE; PS00092; TSP1; 5.

DR PROSITE; PS00142; ZINC_PROTEASE; 1.

KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen.

KW Repeat; Extracellular matrix.		POTENTIAL.
FT SIGNAL	1 27	
FT PROPEP	28 223	BY SIMILARITY.
FT CHAIN	224 1095	ADAMTS-17.
FT DOMAIN	224 452	METALLOPROTEASE.
FT DOMAIN	453 542	DISINTEGRIN-LIKE.
FT DOMAIN	543 598	TSP TYPE-1 1.
FT DOMAIN	599 701	CYS-RICH.
FT DOMAIN	702 779	SPACER.
FT DOMAIN	800 860	TSP TYPE-1 2.
FT DOMAIN	861 922	TSP TYPE-1 3.
FT DOMAIN	925 968	TSP TYPE-1 4.
FT DOMAIN	972 1029	TSP TYPE-1 5.
FT DOMAIN	1045 1084	PLAC.
FT DOMAIN	60 120	ARG-RICH.
FT SITE	201 201	CYSTEINE SWITCH (POTENTIAL).
FT METAL	389 389	ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE	390 390	BY SIMILARITY.
FT METAL	393 393	ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL	399 399	ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD	167 167	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	483 483	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	785 785	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	790 790	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	832 832	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	839 839	N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD	894 894	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE	1095 AA; 121099 MW; A5C6540484754D5D CRC64;	

Query Match		24.0%; Score 274; DB 1; Length 1095;
Best Local Similarity		31.2%; Pred. No. 1.1e-15;
Matches		69; Conservative 36; Mismatches 98; Indels 18; Gaps 6;

QY	8	LELLVAVGPDVFOAH-QEDTERVLTNINIGALLRDPISLGAQPRVHLVKMVLTEPEGA 66
Db	234	VELIWDADADVQYHGAEAAQRFILTVNMVYMFQHSLSGIKINIQIKLVLLRQRPAX 293
QY	67	PNITANLTSSILSYCGW-----SQINPEDDTPGHADLVLYITREDEL-PDG 114
Db	294	LSIGHHGERSLESFCHWQNEEYGGARYLGNQVPGKDDPPLVDAAVFVTRTDFCVHKDE 353
QY	115	NRQVRGVTLQGGAGSPWSCSLITEDTGFDLGVITIAHEIGHSFGLBHDGAPGCGPFGHV 174
Db	354	PCDTVGIAYLGGVCSAKKCVLAEDNGLNLAFTIAHELGHNLGMNDD-DHSSCAGRSHI 412
QY	175	MASDGAAPR--AGLAWSPCSRRLSLISAGRARC--VNDP 211
Db	413	MSGEWKGRNPSDLSSWSSCRDLENFLKSKVSTCLLVTD 453

Search completed: March 13, 2004, 07:39:53
Job time : 4.45833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 21.1771 Seconds
(without alignments)
3188.394 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311

Perfect score: 1141

Sequence: 1 AAGGTLHLELLVAVGPDVDFQ.....QLLSLSAGRCVWDPPRP 214

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phage.*

10: sp_plant.*

11: sp_rodent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriaph.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1141	100.0	1427	4 Q96L37	Q96L37 homo sapien
2	347	30.4	1009	11 Q8BKX1	Q8BKX1 mus musculus
3	347	30.4	1600	11 Q8L1B3	Q8L1B3 mus musculus
4	340	29.8	769	5 Q8MRU5	Q8MRU5 drosophila
5	340	29.8	1059	5 Q9W493	Q9W493 drosophila
6	307.5	27.0	1070	11 Q8CG28	Q8CG28 mus musculus
7	304.5	26.7	967	4 Q8NE26	Q8NE26 homo sapien
8	303.5	26.6	759	6 Q8H2M8	Q8H2M8 equus caball
9	298.5	26.2	900	11 Q8K206	Q8K206 mus musculus
10	295	25.9	2165	5 Q19791	Q19791 caenorhabdi
11	291	25.5	839	6 Q7YS95	Q7YS95 bos taurus
12	286	25.1	833	11 Q8K384	Q8K384 mus musculus
13	286	25.1	845	11 Q8BNJ2	Q8BNJ2 mus musculus
14	283.5	24.8	1092	11 Q8BKAL	Q8BKAL mus musculus
15	263.5	23.1	1091	5 Q3W1Z6	Q3W1Z6 drosophila
16	261.5	22.9	497	11 Q8BZD1	Q8BZD1 mus musculus

17	253.5	22.2	790	5 Q8T458	Q8T458 drosophila
18	253.5	22.2	790	5 Q8NWE1	Q8NWE1 drosophila
19	224.5	19.7	623	11 -Q8BGP4	Q8BGP4 mus musculu
20	212	18.6	1686	5 Q8XBO	Q8XBO drosophila
21	174.5	15.3	872	5 Q22580	Q22580 caenorhabdi
22	168.5	14.8	1461	5 Q8MRA8	Q8MRA8 caenorhabdi
23	163.5	14.3	192	6 Q95N24	Q95N24 equus caball
24	163	14.3	448	11 Q91VS9	Q91VS9 mus musculu
25	162	14.2	203	11 Q8CH80	Q8CH80 cavia porce
26	160	14.0	809	11 Q8CA82	Q8CA82 mus musculu
27	160	14.0	864	11 Q8C720	Q8C720 mus musculu
28	159	13.9	269	6 Q9GL54	Q9GL54 cryptotlagus
29	153.5	13.5	975	5 Q7Z291	Q7Z291 caenorhabdi
30	153.5	13.5	1020	5 Q8IU50	Q8IU50 caenorhabdi
31	153.5	13.5	1020	5 Q19204	Q19204 caenorhabdi
32	150	13.1	561	4 Q86XK9	Q86XK9 homo sapien
33	146	12.8	482	13 Q9FVK9	Q9FVK9 agkistrodon
34	146	12.8	509	5 Q20930	Q20930 caenorhabdi
35	139	12.2	914	13 Q12960	Q12960 xenopus lae
36	135.5	11.9	617	13 Q90499	Q90499 echis pyram
37	134	11.7	339	4 Q8NEK2	Q8NEK2 homo sapien
38	133.5	11.7	812	4 Q8NDW6	Q8NDW6 homo sapien
39	130	11.4	936	5 Q86F79	Q86F79 drosophila
40	130	11.4	959	5 Q9V7H0	Q9V7H0 drosophila
41	128.5	11.3	480	13 Q90220	Q90220 agkistrodon
42	128	11.2	476	5 Q8MT72	Q8MT72 drosophila
43	128	11.2	761	11 Q8CDV3	Q8CDV3 mus musculu
44	127	11.1	610	13 Q8AWI5	Q8AWI5 agkistrodon
45	126	11.0	243	13 Q7TIT3	Q7TIT3 bothrops ja

ALIGNMENTS

RESULT 1

Q96L37 PRELIMINARY; PRT; 1427 AA.

AC Q96L37
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUS=Liver;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,
Fujikawa K.,
RT "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a
Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura.";
RL J. Biol. Chem. 276:41059-41063(2001).
DR EMBL; AY055376; AAL17652.1; -
DR MEROPS; M12.241; -

DR Genew; HGNC:1366; ADAMTS13.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept M_Zn_BS.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00209; TSPL; 7.
DR PROSITE; PS0215; ADAM MEPRO; 1.
DR PROSITE; PS00392; TSPL; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Signal.
FT SIGNAL 1 33 POTENTIAL.

```
FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB1BC3AAB1A4442 CRC64;

Query Match 100.0%; Score 1141; DB 4; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.1e-94;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGIIHLELLVAVGPDVFOAHQEDTERRYVLTNIGAEILRDSLGAFVHLVKWVIL 60
DB 75 AAGGIIHLELLVAVGPDVFOAHQEDTERRYVLTNIGAEILRDSLGAFVHLVKWVIL 134
QY 61 TPEGAPNITANITSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLEPDGNGVRG 120
DB 135 TPEGAPNITANITSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLEPDGNGVRG 194
QY 121 VTQLGACSPWSCSLITDGTDFGLGVTTIAHEIGHSGFLEHGDGAPGCGPSCGHVNASDGA 180
DB 195 VTQLGACSPWSCSLITDGTDFGLGVTTIAHEIGHSGFLEHGDGAPGCGPSCGHVNASDGA 254
QY 181 APRAGLAWSPCSRRLSLLSAGRARCVWDPPRP 214
DB 255 APRAGLAWSPCSRRLSLLSAGRARCVWDPPRP 288

RESULT 2
Q8BK1 PRELIMINARY; PRT; 1009 AA.
AC Q8BK1;
DT 01-VAR-2003 (TrEMBLrel. 23, Created)
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)
DE ADAMTS-12 precursor (Fragment).
GN A1605170.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs".
RL Nature 420:563-573(2002).
DR EMBL; AK048612; BAC33391.1; -.
DR F1R; PT0546; PT0698.
DR MGD; MGI:2146046; A1605170.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprylisin; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SMC0209; TSP1; 4.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON TER 1009
SQ SEQUENCE 1009 AA; 113352 MW; 7B63218CFF50FDB1 CRC64;

Query Match 30.4%; Score 347; DB 11; Length 1009;
Best Local Similarity 36.5%; Pred. No. 1.2e-22;
Matches 77; Conservative 36; Mismatches 90; Indels 8; Gaps 5;

QY 8 LELLVAVGPDVFOAH-QEDTERRYVLTNIGAEILRDSLGAFVHLVKWVILTEPEGA 66
DB 252 VETLVVADTKVEYHGSENVESYILTMNMTGLFHSPSIGNLVHVVVRLILLEEEQ 311
QY 67 PNITANLTSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLEPDGNGR--QVRGVTL 124
DB 312 LKIVHAEKTLSSFCWKQKSIKPSDLNPNVHDAVLITRKDI-CAGVNRPCETLGLSQL 370
QY 125 GCACSPWSCSLITDGTDFGLGVTTIAHEIGHSGFLEHGDGAPGCGPSCG--HVMASDGA 181
DB 371 SGMCQPHRSCNINEDSGLPFLAFTIAHELHSGFGLQHDGKE--NDCEPVGRRHPYIMSQQIQY 429
QY 182 PRAGLAWSPCSRRLSLLSAGRARCVWDPP 212
DB 430 DTFPLTWSKSKSEYITRFLDRGEGFCLDDIP 460

RESULT 3
Q811B3 PRELIMINARY; PRT; 1600 AA.
AC Q811B3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloprotease disintegrin 12 protein.
GN ADAMTS-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Cal S., Lopez-Otin C.;
RT "Mouse ADAMTS-12";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ537452; CAD60967.1; -.
DR F1R; PT0546; PT0698.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR008084; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprylisin; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SMC0209; TSP1; 8.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Integrin; Protease.
SQ SEQUENCE 1600 AA; 177791 MW; 083089D356E38C5D CRC64;

Query Match 30.4%; Score 347; DB 11; Length 1600;
Best Local Similarity 36.5%; Pred. No. 2.2e-22;
Matches 77; Conservative 36; Mismatches 90; Indels 8; Gaps 5;

QY 8 LELLVAVGPDVFOAH-QEDTERRYVLTNIGAEILRDSLGAFVHLVKWVILTEPEGA 66
DB 252 VETLVVADTKVEYHGSENVESYILTMNMTGLFHSPSIGNLVHVVVRLILLEEEQ 311
QY 67 PNITANLTSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLEPDGNGR--QVRGVTL 124
DB 312 LKIVHAEKTLSSFCWKQKSIKPSDLNPNVHDAVLITRKDI-CAGVNRPCETLGLSQL 370
QY 125 GCACSPWSCSLITDGTDFGLGVTTIAHEIGHSGFLEHGDGAPGCGPSCG--HVMASDGA 181
DB 371 SGMCQPHRSCNINEDSGLPFLAFTIAHELHSGFGLQHDGKE--NDCEPVGRRHPYIMSQQIQY 429
QY 182 PRAGLAWSPCSRRLSLLSAGRARCVWDPP 212
```

Db	430	DTPLTWSKSEYITRFLDRGRGFLDDIP	460
RESULT 4			
Q8MR15			
ID	Q8MR15	PRELIMINARY;	PRT; 769 AA.
AC	Q8MR15		
OC	Q8MR15		
DT	01-OCT-2002 (TReMBLrel. 22, Created)		
DT	01-OCT-2002 (TReMBLrel. 22, Last sequence update)		
DT	01-OCT-2003 (TReMBLrel. 25, Last annotation update)		
DE	GH22104p.		
GN	CG4096.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Berkeley;		
RX	MEDLINE=20196006; PubMed=10731132;		
RA	Adams M.D., Celnikner S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,		
RA	April J.F., Agbayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,		
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,		
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,		
RA	Burtis K.C., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
RA	Cherry J.M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,		
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,		
RA	Hoskin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
RA	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
RA	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,		
RA	Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
RA	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
RA	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,		
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,		
RA	Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,		
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,		
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,		
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,		
RT	"The genome sequence of Drosophila melanogaster."		
RL	Science 287:2185-2195(2000).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,		
RA	Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,		
RA	Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,		
RA	Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,		
RA	Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,		
RA	Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,		
RA	Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,		
RA	Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,		
RA	McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,		
RA	Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,		
RA	Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,		
RA	Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,		
RA	Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.,		
RT	"Sequencing of Drosophila melanogaster genome."		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBSJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,		
RA	Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,		

Query Match	29.8%	Score 340;	DB 5;	Length 769;
Best Local Similarity	35.0%	Pred. No. 3.7e-22;		
Matches	75;	Conservative 31;	Mismatches 99;	Indels 10;
Gaps	5;			
QY	7	HLELLVAVGPDVFAHQEDTERRYVLTNLNIGALLRDPISLGAQFRVHLVXWVILPEPG-	65	
Db	314	HVTLLI-VADATWSAFRLDNGVLLTNNVSYALKYDPSIGNSTHVVVRIQLDEESQ	372	
QY	66	-APNITANLSSLLSVGWSQTINPDDTDFGHADLVLYITRFDLDPGNRQVGTQL	124	
Db	373	LQLNLICNAQKNLDRFCQWQKLNKSGKDPHHQVAILTR--KNICANNCTMLGLANV	430	
QY	125	GGACSPWCSCLITDTDFGLVTTAHEIGHSGFLEHDCAPGSGC-----GPSGHVMSDGA	180	
Db	431	GGMKPKQSCSVNEDNGIMLSHTITHELGHFGFMFHDFTAK-IGCHPRVGPVHIIMTPTFG	489	
QY	181	APRAGLAWSCSRQLLSLSAGRARCVDPRP	214	
Db	490	ADTLQVCWNCSSRKYYTHFLDQGLGECLEDDPTPT	523	
RESULT 5				
Q9W493				
ID	Q9W493	PRELIMINARY;	PRT; 1059 AA.	
AC	Q9W493;			

Tupy J.L., Bergman C., Berman E., Carlson J.W., Celniker S.E., Clamp M., Drysdale R., Emmert D., Friese E., de Grey A., Harris N., Krommiller B., Marshall B., Millburn G., Richter J., Russo S., Seale S.M.J., Smith B., Shu S., Smutniak F., Whitfield E., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.; "Annotation of Drosophila melanogaster genome."; Submitted (MAR-2000) to the EMBL/GenBank/DDJB databases.

RN [4]
RN
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
R submitted (MAR-2000) to the EMBL/GenBank/DDJB databases.
RL
RN SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DDJB databases.
DR EMBL; AF603435; AAFA6065.2; -
DR HSSP; P15167; IATL.
DR MEROPS; M12.231; -.
DR FlyBase; FBgn0029791; CG4096.
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0008270; Zinc ion binding; IEA.
DR GO; GO:0006508; Proteolysis and peptidolysis; IEA.
DR InterPro; IPRO01590; Peptide_M12B_N.
DR InterPro; IPRO02870; Peptide_M12B_N.
DR InterPro; IPRO06025; Pept_M_Zn_BS.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Replicysin; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 5.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSEQUENCE 1059 AA; 119195 MW; AD3C99E47618F3D7 CRC64;

Query Match 29.8%; Score 340; DB 5; Length 1059;
Best Local Similarity 35.0%; Pred.No. 5.6e-22;
Matches 75; Conservative 31; Mismatches 99; Indels 10; Gaps 5;

Qy 7 HLELLVAVGPDVFQAHQEDTERTYVLTNIGAEILRDPSLGACFRHVKMILTPEPG- 65
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 331 HVETLI-VADATMSAFHRDLNGYLITMMVMYSALYKDPDSIGNSIHVVRIIQLDEESQ 389
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 66 APNITANLTSSLVSVCWGSQTINPEDDTDPGHADVLYITFDLPDGNRVGVTVQL 124
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 390 LQLNLTLNQAKNLDFRCWSQHKLKNKGSKDPHHHDVAIITR--KNICANNMTGLNAV 447
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 125 GGACSTWCSCITTDGTGDVGTTIAHTCHSGFLSHDGAPGGGC---GPSGHVVASDGA 180
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 448 GMCKPKQSQCSDNEINGMLSHTHIEUHGFGMHPTAK-IQGPRVGPVIHMTPFTG 506
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 181 APRAGLAWSCRRSLLSLAGRARCVWDPPRP 214
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 507 ADTIQVCSNSCRKYITHFDLQGLCECLDDPTTP 540
Db :|||:::|||:::|||:::|||:::|||:::|||:

RESULT 6

OBCG28 PRELIMINARY; PRT; 1070 AA.

ID OBCG28;
DC AC
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Zinc metalloendopeptidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurogathi; Muridae; Murinae; Mus.
OX NCBI_taxID=10990;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;

Bratwaite M., Nagaraja R., Abe K.; "Genomic Sequence Analysis in the Mouse t-complex Region.";
Submitted (JUL-2002) to the EMBL/GenBank/DDJB databases.
EMBL; AF528163; AAC17380.1; -
GO; GO:0004222; F-metalloendopeptidase activity; IEA.
GO; GO:0008270; Zinc ion binding; IEA.
GO; GO:0006508; Proteolysis and peptidolysis; IEA.
InterPro; IPRO01590; Peptide_M12B_N.
InterPro; IPRO02870; Peptide_M12B_N.
InterPro; IPRO06025; Pept_M_Zn_BS.
InterPro; IPRO00884; TSP1.
InterPro; IPRO08085; TSP1.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Replicysin; 1.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1; 5.
PROSITE; PS50215; ADAM_MEPRO; 1.
PROSITE; PS50092; TSP1; 5.
PROSITE; PS00142; ZINC_PROTEASE; 1.
PROSEQUENCE 1070 AA; 117203 NW; 6B896425EBA2B8D8 CRC64;

Query Match 27.0%; Score 307.5; DB 11; Length 1070;
Best Local Similarity 34.8%; Pred.No. 5.1e-19;
Matches 78; Conservative 34; Mismatches 91; Indels 21; Gaps 7

Qy 7 HLELLVAVGPDVFQAHOEDTERTYVLTNIGAEILRDPSLGACFRHVKMILTPEPG 65
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 206 YVTLVVADKMWAYHGQRVEQYVALMNIVAKLFQSSLGNTVINLVTRILLITEDQP 265
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 66 APNITANLTSSLVSVCWGSQTI-----NPEDDTDPHADVLYITFPDL----ELPDG 114
Db :|||:::|||:::|||:::|||:::|||:::|||:
Qy 266 TLEITHHAGKLDSECFKWQKSIVSHSGNAIPENGVANHTAVITRYDICIKNKPCG 325
DB 115 NRQVRGYTOLGGACSTWCSCITTEDTGVDLGMTTAHICHSGFLSHDHGAPGGCGSPSHG- 173
DB 326 ---TLGLAPVGMCMERERSCSINEDIGLATFTAHTHAIGHTFGNMHDGV-GNGCGARGOD 381
Qy 174 ---VNASGAAPRAGLAWSCRRSLLSLAGRARCVWD-PPR 213
DB :|||:::|||:::|||:::|||:::|||:::|||:
Qy 382 PAKLMAAHIKTMPFWVSWSCSRDYITSFLDSLGLCLNNRPRE 425
DB :|||:::|||:::|||:::|||:::|||:::|||:

RESULT 7

OBCG26 PRELIMINARY; PRT; 967 AA.

ID QBCG26;
DC AC
DT 01-OCT-2002 (TREMBLrel. 22, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)

DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiadae; Homo.
OX NCBI_TaxID=9606;
[1]
RN RP SEQUENCE FROM N.A.
RC TIISUB-Testis;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DDJB databases.
DR EMBL; BC036515; AAAH36515.1; -
DR GO; GO:0004222; F-metalloendopeptidase activity; IEA.
DR GO; GO:0008270; Zinc ion binding; IEA.
DR GO; GO:0006508; Proteolysis and peptidolysis; IEA.
DR InterPro; IPRO06586; ADAM cysteine.
DR InterPro; IPRO01128; Cytochrome_P450.
DR InterPro; IPRO01590; Peptide_M12B_N.
DR InterPro; IPRO02870; Peptide_M12B_N.
DR InterPro; IPRO06025; Pept_M_Zn_BS.
DR InterPro; IPRO00884; TSP1.
DR InterPro; IPRO08085; TSP1.
PFam; PF01562; Pep_M12B_propep; 1.

Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS02015; ADAM_MEPRO; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 967 AA; 105387 MW; FF1D399674201C3D CRC64;

Query Match 26.7%; Score 304.5; DB 4; Length 967;
Best Local Similarity 29.8%; Pred. No. 8.3e-19;
Matches 65; Conservative 40; Mismatches 19; Gaps 5

OY 7 HLELLVAVGPDVFQAHQEDTERTYVLTNLNICGAELRDPGLGAQFRVHLVKWVILTEPEGA 66
:
Db 259 YVETMLVDQMMAEFHSGSLGHYLLTLFSVAARLYKHPSIRNSVSVLVVKILVIHQKG 318
:
OY 67 PNITANTLSLLSCVGWSQTINPEDDTDPGHADLVLIYTRFDLPDGNR--QVRGVQL 124
:
Db 319 PEVTSNAALTNRNFCNQKQHNPSSDRDAEYHTAILFTRODL---CGSQTCDTLGMADV 375
:
OY 125 GGACSPWSCIIVTDGTGDLGVTTAHEIGHGSFGLEHDGAPG-----SGCGPSSHVMAS---- 177
:
Db 376 GTVCDPSPRSCTVIDDDQLQAFTTAHELGHGFVNPHFDDAKQCASLNGYNQDSHMMSMLS 435
:
OY 178 --DGAAAPRAGLAWPCSRRLROLLSLSAGRARCVCWDPPR 213
:
Db 436 NLDSQP-----WSPCSAYMITSFLDNHGHCMDKPQ 468
: :

RESULT 8
Q8HZM8 PRELIMINARY; PRT; 759 AA.

ID Q8HZM8 AC Q8HZM8
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloproteinase (Fragment).
GN ADAMTS-1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Boerboom D., Russell D.L., Richards J.S., Sirois J.;
RT "Coordinated Regulation of Transcripts Encoding ADAMTS-1 and
RT Progesterone Receptor in Equine Preovulatory Follicles.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF541975; AANL7331.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pspt_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON TRX
SQ SEQUENCE 759 AA; 83410 MW; C18E4C048918CB9 CRC64;

Query Match 26.6%; Score 303.5; DB 6; Length 759;

[illegible]

463 RGVSVSWSCSRQYHKKFLSTAAQICLADQPKP 495

DB ID Q19791 PRELIMINARY; PRT; 2165 AA.

AC Q19791; Q27524;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE F25H8.3 protein.

GN F25H8.3

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;

OC Rhabditiidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN RN

RP SEQUENCE FROM N.A.

RA Gajadasty S.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton A., Dear S., Du Z., Durbin R., Favellio A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R., Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinso-Spratt J., Wohlman P.;

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RT Nature 368:32-38(1994).

RL DB 1

QY Query Match 25.9%; Score 295; DB 5; Length 2165;

Best Local Similarity 30.4%; Pred. No. 1.7e-17;

Matches 70; Conservative 30; Mismatches 104; Indels 26; Gaps 3;

DB ID Q19791 PRELIMINARY; PRT; 2165 AA.

AC Q19791; Q27524;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE F25H8.3 protein.

GN F25H8.3

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;

OC Rhabditiidae; Peloderinae; Caenorhabditis.

OX NCBI_TaxID=6239;

RN RN

RP SEQUENCE FROM N.A.

RA Gajadasty S.;

RL Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=94150718; PubMed=7906398;

RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A., Craxton A., Dear S., Du Z., Durbin R., Favellio A., Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L., Jones M., Kershaw J., Kirsten J., Laister N., Latreille P., Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M., Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkneen R., Smaldon N., Smith A., Sonhammer E., Staden R., Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R., Watson A., Weinstock L., Wilkinso-Spratt J., Wohlman P.;

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";

RT Nature 368:32-38(1994).

RL DB 1

QY Query Match 25.9%; Score 295; DB 5; Length 2165;

Best Local Similarity 30.4%; Pred. No. 1.7e-17;

Matches 70; Conservative 30; Mismatches 104; Indels 26; Gaps 3;

[illegible]

QY 64 EGAPNITANLTSSLVCGWSQTINPEDDTPDGHADLVLYITRFDL-ELPDGNRQVRGVT 122
Db 181 EGL-LINHADGSLNSFCQWSALVKNKGRHDA---ILLTGDCSWKNEPCDTLGF 236
QY 123 QUGGACSPWTSCLITDTPDGLGVTIAHGHGFLHGDGAPGCGPS-GHVNASDGA 181
Db 237 PISGWCYSKYRSCNTINEDTGLGFLTAHGHGFLHGDGAPGCGPS-GHVNASDGA 295
QY 182 PRAGLAWSPCSRQLLSLSSAGRARCVDPPR 213
Db 296 NNGVFSWSSCSROYLKKFLSTPQAGCLVDBPK 327
RESULT 15
Q9W126 PRELIMINARY; PRT; 1091 AA.
AC Q9W126;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG3622 protein.
GN CG3622 OR BCDNA:GM08694.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Botchava D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferrar C., Ferrier S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR ENBL; AE003458; AAF46905.1; --
DR FlyBase; FBgn0034778; CG3622.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.

DR GO; GO:0009270; F:zinc ion binding; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_zn_BS.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
DR PROSITE; PS0215; ADAM MEPRO; 1.
DR PROSITE; PS0092; TSP1; 1.
DR PROSITE; PS00142; ZINC PROTEASE; 1.
SQ SEQUENCE 1091 AA; 58967382F12565BE CRC64;
Query Match 23.1%; Score 263.5; DB 5; Length 1091;
Best Local Similarity 32.7%; Pred. No. 5.1e-15;
Matches 73; Conservative 38; Mismatches 85; Indels 27; Gaps 11;
QY 6 LHLELLVAVGPDVFOAHQE-----DTE-----RVLTNLNIGALL-RDPSLGAQPRVHLV 55
Db 286 LFIETAFVDSLY-AHMQKQFTNTESKVSFLLAMIN-GVQLYHHTLORRINFVLK 343
QY 56 KXVILT--EPEGAPNITANLTSSLVCGWSQTINPEDDTPDGHADLVLYITRFDLELPD 113
Db 344 RLEIWSWDPPLVLR-SRDVENYLSFCKWQKLNPFSDADPLHYDHALVLTGLDLVYD 402
QY 114 ---GNRQVRGVTOLGGACSPWTSCLITDTPDGLGVTIAHGHGFLHGDGAPGCGCP 170
Db 403 KGKANSQVVGMAIVKGMCTSIYCTINEAKHFESVFFVAHEIGHNLGVRHD-AKEISCDP 461
QY 171 SGHVMASDGAAPRAG---LAWSPCSRQLLSLSSAGRARCVD 210
Db 462 TWHIM-----SPKLGSGKVTWSKSRTYLEDFLMDPOAECLFD 499

Search completed: March 13, 2004, 07:43:15
Job time : 23.1771 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 31.3198 Seconds
(without alignments)
1930.574 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311

Perfect score: 1141

Sequence: 1 AAGGILHLELLVAVGPDVFQ.....QLLSLSAGRCVWDPPRP 214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 29Jan04:*

1: Geneseqp1980s.*

2: Geneseqp1980s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1141	100.0	933	5	Aau79217 Human ADA
2	1141	100.0	1297	6	Aao16618 Human von
3	1141	100.0	1353	5	Aae24449 Human von
4	1141	100.0	1353	6	Aao16617 Human von
5	1141	100.0	1415	5	Abb04153 Human ADA
6	1141	100.0	1427	5	Aae24450 Human von
7	1141	100.0	1427	7	Add94038 Human agg
8	1141	100.0	1445	5	Abb98125 Human PMM
9	1065	93.3	312	6	Aao16621 Human von
10	1065	93.3	1323	6	Aao16620 Human von
11	1065	93.3	1378	6	Aao16619 Human von
12	1059.5	92.9	270	6	Aao16622 Human von
13	1057	92.6	242	5	Aau97641 Human agg
14	1057	92.6	242	7	Add94031 Human agg
15	1046	91.7	203	4	Aag63826 Amino aci
16	1046	91.7	203	6	Abg73948 Metallopr
17	1046	91.7	1120	4	Aag63829 Amino aci
18	1046	91.7	1120	6	Abg74113 Human MOV
19	1029.5	90.2	984	6	Abg96171 Human NAT
20	969	84.9	184	6	Aao16614 Human von
21	843	73.9	161	6	Aao16612 Human von
22	843	73.9	161	6	Aao16623 Human von
23	773	67.7	148	5	Aae24448 Human von
24	613	53.7	118	3	AAB42668 Human ORF
25	361	31.6	1686	4	Aae00934 Human 278

ALIGNMENTS

RESULT 1

AAU79217

ID AAU79217 standard; protein; 933 AA.

XX AC AAU79217;

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Aae00913 Human 278
Aab74944 Human ADA
Aab41379 Human ORF
Aab72283 Human ADA
Aab86949 Human met
Abb60410 Drosophil
Aab96136 Human NOV
Ada50466 Human pro
Add98130 Human PMM
Add44592 Polypept1
Aau72897 Human met
Aau08955 Human ADA
Aao3581 Human pro
Aao3582 Human pro
Aao3576 Human pro
Aao3577 Human pro
Aao3583 Human pro
Aao3572 Human pro
Aab86947 Human met
Abg76502 DNA encod

CC kinds of secondary arterial hypertension), and peripheral vascular
 CC diseases (e.g., chronic peripheral arterial occlusive disease, acute
 CC arterial thrombosis and embolism, inflammatory vascular disorders, etc),
 CC liver disorders and chronic obstructive pulmonary disease. The sequences
 CC are useful in diagnostic assays for detecting diseases and abnormalities
 CC or susceptibility to diseases and abnormalities related to the presence
 CC of mutations in the nucleic acid sequences which encode the protein. The
 CC sequences are also useful for modulating ADAM-TS-like protein activity in
 CC a disease condition. This sequence represents the human ADAM-TS-like
 CC protein
 XX
 SQ Sequence 933 AA;
 Query Match 100.0%; Score 1141; DB 5; Length 933;
 Best Local Similarity 100.0%; Pred. No. 6.6e-110;
 Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDSLGAFVHLVKMVL 60
 DB 78 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDSLGAFVHLVKMVL 137
 QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 120
 DB 138 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 197
 QY 121 VTQLGACPTWSCLITETGFDLGVTTIAHEIGHSGFLEHDPGAPSGCGPSGHVMSDGA 180
 DB 198 VTQLGACPTWSCLITETGFDLGVTTIAHEIGHSGFLEHDPGAPSGCGPSGHVMSDGA 257
 QY 181 APRAGLAWSPCSRRLSLLSAGRARCWDPDPRP 214
 DB 258 APRAGLAWSPCSRRLSLLSAGRARCWDPDPRP 291

RESULT 2
 AAO16618
 ID AAO16618 standard; protein; 1297 AA.
 XX
 AC AAO16618;
 XX
 DT 15-MAY-2003 (first entry)
 XX
 DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #5.
 XX
 KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
 KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
 KW myocardial infarction; cerebral infarction; arteriosclerosis;
 KW platelet thrombosis; stenosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200288366-A1.
 XX
 PD 07-NOV-2002.
 XX
 PF 25-APR-2002; 2002WO-JP004141.
 XX
 PR 25-APR-2001; 2001JP-00128342.
 PR 27-JUL-2001; 2001JP-00227510.
 PR 28-SEP-2001; 2001JP-00302977.
 PR 23-JAN-2002; 2002JP-00017596.
 XX
 (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
 PA
 SOejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
 WPI; 2003-120479/11.
 DR N-PSDB; ABT32584.
 XX
 PT von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
 PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
 PT developing drugs for e.g. myocardial infarction and cerebral infarction.
 XX

PS Claim 4; Page 92-101; 144pp; Japanese.
 CC The invention comprises the amino acid and coding sequence of a von
 CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
 CC the invention are useful in the diagnosis and treatment of
 CC thrombocytopenic purpura, and in developing drugs for myocardial
 CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
 CC and stenosis. The present amino acid sequence represents a human von
 CC Willebrand factor (vWF)-cleaving enzyme-related protein
 XX
 SQ Sequence 1297 AA;
 Query Match 100.0%; Score 1141; DB 6; Length 1297;
 Best Local Similarity 100.0%; Pred. No. 1.1e-109;
 Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDSLGAFVHLVKMVL 60
 DB 1 AAGGILHLELLVAVGPDVFOAHQEDTERVYLTNLNIGAEILLRDSLGAFVHLVKMVL 60
 QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 120
 DB 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 120
 QY 121 VTQLGACPTWSCLITETGFDLGVTTIAHEIGHSGFLEHDPGAPSGCGPSGHVMSDGA 180
 DB 121 VTQLGACPTWSCLITETGFDLGVTTIAHEIGHSGFLEHDPGAPSGCGPSGHVMSDGA 180
 QY 181 APRAGLAWSPCSRRLSLLSAGRARCWDPDPRP 214
 DB 181 APRAGLAWSPCSRRLSLLSAGRARCWDPDPRP 214

RESULT 3
 AAE24449
 ID AAE24449 standard; protein; 1353 AA.
 XX
 AC AAE24449;
 XX
 DT 04-OCT-2002 (first entry)
 XX
 DE Human Von Willebrand factor-cleaving protease fragment #2.
 XX
 KW Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;
 KW transgenic animal; immunisation; thromboembolic disease; preclampsia;
 KW thrombotic thrombocytopenic purpura; TRP; Henoch-Schonlein purpura;
 KW thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;
 KW transgenic; anticoagulant.
 XX
 OS Homo sapiens.
 XX
 PN WO200242441-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 20-NOV-2001; 2001WO-EP013391.
 XX
 PR 22-NOV-2000; 2000US-00721254.
 PR 12-APR-2001; 2001US-00833328.
 XX
 PA (BAXT) BAXTER AG.
 XX
 PI Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
 PI Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
 PI Zimmermann K, Voelkel D;
 XX
 DR WPI; 2002-479950/51.
 XX
 PT Novel isolated or substantially purified Von Willebrand factor-cleaving
 PT protease, useful for producing preparation for therapy of thrombosis and
 PT thromboembolic disease such as thrombotic thrombocytopenic purpura.
 XX
 PS Claim 1; Page 64-68; 93pp; English.

XX The invention relates to an isolated or substantially pure von Willebrand
CC factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
CC purifying vWF which involves providing vWF-cp as a ligand, contacting a
CC solution comprising vWF with the polypeptide ligand under conditions
CC where vWF is bound to the ligand and recovering from the ligand purified
CC vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
CC which involves immunising an animal with vWF-cp and isolating the anti-
CC vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
CC producing a preparation of prophylaxis and therapy of thrombosis and
CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
CC Henoch-Schönlein purpura, pre-eclampsia, neonatal thrombocytopenia or
CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing
CC plasmatic or recombinantly produced vWF. The invention is useful for
CC construction expression systems and generating transgenic animals which
CC express the polypeptide in vivo. The present sequence is human vWF-cp
CC fragment
XX
SQ Sequence 1353 AA;
Query Match 100.0%; Score 1141; DB 5; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.1e-109; Indels 0; Gaps 0;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLELLVAVGPDVFOAHQEDTERYVLTNINIGAEILLRDPGLGAQFRVHLVKWVIL 60
DB 1 AAGGILHLELLVAVGPDVFOAHQEDTERYVLTNINIGAEILLRDPGLGAQFRVHLVKWVIL 60
QY 61 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNRQVRG 120
DB 61 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNRQVRG 120
QY 121 VTOLGACSTWSCLITDFTGLGVITAHIGHSFGLHGDGAPGCGPSGHVNASDGA 180
DB 121 VTOLGACSTWSCLITDFTGLGVITAHIGHSFGLHGDGAPGCGPSGHVNASDGA 180
QY 181 APRAGLAWSPCSRRLSLLSAGRARCVDPPRP 214
DB 181 APRAGLAWSPCSRRLSLLSAGRARCVDPPRP 214
RESULT 4
AAO16617
ID AAO16617 standard; protein; 1353 AA.
XX AAO16617;
AC AAO16617;
XX 15-MAY-2003 (first entry)
DT Human von Willebrand factor (vWF)-cleaving enzyme-related protein #4.
XX Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
KW myocardial infarction; cerebral infarction; arteriosclerosis;
KW platelet thrombosis; stenosis.
XX Homo sapiens.
OS WO200288366-A1.
PN 07-NOV-2002.
PD 25-APR-2002; 2002WO-JP004141.
XX 25-APR-2001; 2001JP-00128342.
PR 27-JUL-2001; 2001JP-00227510.
PR 28-SEP-2001; 2001JP-00302977.
PR 25-JAN-2002; 2002JP-00017596.
XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
PA Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
XX

DR WPI; 2003-120479/11.
DR N-PSDB; ABT32583.
XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
PT developing drugs for e.g. myocardial infarction and cerebral infarction.
XX Claim 4; Page 82-92; 144pp; Japanese.
PS The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein
XX
SQ Sequence 1353 AA;
Query Match 100.0%; Score 1141; DB 6; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.1e-109; Indels 0; Gaps 0;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLELLVAVGPDVFOAHQEDTERYVLTNINIGAEILLRDPGLGAQFRVHLVKWVIL 60
DB 1 AAGGILHLELLVAVGPDVFOAHQEDTERYVLTNINIGAEILLRDPGLGAQFRVHLVKWVIL 60
QY 61 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNRQVRG 120
DB 61 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNRQVRG 120
QY 121 VTOLGACSTWSCLITDFTGLGVITAHIGHSFGLHGDGAPGCGPSGHVNASDGA 180
DB 121 VTOLGACSTWSCLITDFTGLGVITAHIGHSFGLHGDGAPGCGPSGHVNASDGA 180
QY 181 APRAGLAWSPCSRRLSLLSAGRARCVDPPRP 214
DB 181 APRAGLAWSPCSRRLSLLSAGRARCVDPPRP 214
RESULT 5
ABB04153
ID ABB04153 standard; protein; 1416 AA.
XX ABB04153;
AC ABB04153;
XX 26-MAR-2002 (first entry)
DT Human ADAMTS-M polypeptide.
DE Osteoarthritis; rheumatoid arthritis; inflammatory bowel disease;
XX Crohn's disease; asthma; Alzheimer's disease; organ transplant rejection;
KW cachexia; allergy; cancer; leukaemia; lymphoma; osteoporosis;
KW atherosclerosis; congestive heart failure; myocardial infarction; stroke;
KW neurodegenerative disease; autoimmune disorder; Huntington's;
KW Parkinson's; migraine; pain; depression; multiple sclerosis; burn;
KW infertility; diabetic shock; gene therapy; ADAMTS-M;
KW A Disintegrin And Metalloprotease; thrombospondin domain.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH Peptide 1..97
FT /label= Prodomain
FT /note= "The mature form of the ADAMTS-M protein is
FT processed by furin cleavage of the prodomain"
FT Cleavage-site 94..97
FT /label= Furin_cleavage_site
FT Protein 98..1416
FT /label= Mature ADAMTS-M protein
FT /note= "The mature form of the ADAMTS-M protein is
FT processed by furin cleavage of the prodomain"
FT Domain 98..311

FT	Domain	/label= Metalloprotease_domain	247..272	
FT	FT	/label= Zinc-binding_motif	324..394	
FT	Domain	/label= Disintegrin_domain	410..473	
FT	Domain	/label= Thrombospondin_submotif	419..424	
FT	Domain	/label= Heparin-binding_domain	1099..1156	
FT	Domain	/label= Thrombospondin_submotif		
XX	EP1152055-A1.			
PN	07-NOV-2001.			
PD	24-APR-2001; 2001EP-00303706.			
XX	27-APR-2000; 2000US-0200040P.			
XX	(PFIZ) PFIZER PROD INC.			
XX	Buckbinder L, Mitchell PG, Wachtmann TS, Walsh RT;			
PI	WPI; 2002-084275/12.			
DR	N-PSDB; ABA02549.			
XX	New polynucleotide, useful in gene therapy, particularly for treating or			
FT	preventing e.g. arthritis, Crohn's disease, Alzheimer's disease and organ			
PT	transplant toxicity and rejection, comprises ADAMTS polynucleotide and			
PT	encoded polypeptide.			
XX	Claim 4; Fig 2; 31pp; English.			
XX	The present sequence represents a ADAMTS protein, designated ADAMTS-M,			
CC	that exhibits the characteristics of the ADAM (A Disintegrin And			
CC	Metalloprotease) family of metalloproteases, and contains a			
CC	thrombospondin domain (TS). The protein is encoded by the cDNA given in			
CC	ABAO2549. The specification describes a newly isolated polynucleotide,			
CC	comprising a nucleotide sequence encoding an ADAMTS-M polypeptide as			
CC	given in the specification, or a metalloproteinase, disintegrin domain,			
CC	prodomain or its thrombospondin submotif. The polynucleotide, polypeptide			
CC	and agent are useful for manufacturing a medicament for treating a			
CC	subject in need of altering activity or expression of ADAMTS-M. The			
CC	polynucleotide, ADAMTS-M polypeptide and agent are useful for			
CC	manufacturing a medicament for treating arthritis (osteoarthritis and			
CC	rheumatoid arthritis), inflammatory bowel disease, Crohn's disease,			
CC	asthma, Alzheimer's disease, organ transplant toxicity and rejection,			
CC	cachexia, allergy, cancer (e.g. solid tumour cancer including colon,			
CC	breast, lung, prostate, brain or haematopoietic malignancies including			
CC	leukaemia and lymphoma), osteoporosis, atherosclerosis, aortic aneurysm,			
CC	congestive heart failure, myocardial infarction, stroke, head trauma,			
CC	spinal cord injury, neurodegenerative disease, autoimmune disorders,			
CC	Huntington's disease, Parkinson's disease, migraine, pain, depression,			
CC	multiple sclerosis, abnormal wound healing, burns, infertility or			
CC	diabetic shock. The polynucleotide and polypeptide are also useful for			
CC	diagnosing the diseases above. The polynucleotide is particularly useful			
CC	in gene therapy for treating the diseases cited above			
XX	Sequence 1416 AA;			
SQ	Query Match 100.0%; Score 1141; DB 5; Length 1416;			
	Best Local Similarity 100.0%; Pred. NO. 1.2e-109;			
	Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 AAGGTLHLELVAVGPDVFOAHQEDTERVLTNLNIGALLRDPISLGAQFRVHLVKMVL 60			
DB	98 AAGGTLHLELVAVGPDVFOAHQEDTERVLTNLNIGALLRDPISLGAQFRVHLVKMVL 157			
QY	61 TEPEGAPNITANTSSLSVCGWSQINPEDDDTPGHADLVLYITRFDLELPQGNQVRG 120			
DB	158 TEPEGAPNITANTSSLSVCGWSQINPEDDDTPGHADLVLYITRFDLELPQGNQVRG 217			
QY	121 VTQLGACSPWTSCLITDTGFDLGVTTIAIEHGHSGLEHDSGAPGSGCGPSGHVMSDGA 180			
DB	218 VTQLGACSPWTSCLITDTGFDLGVTTIAIEHGHSGLEHDSGAPGSGCGPSGHVMSDGA 277			
QY	181 APRAGLAWSPCSRRLQLLSLSAGRAFCVWDPPRP 214			
DB	278 APRAGLAWSPCSRRLQLLSLSAGRAFCVWDPPRP 311			
XX	RESULT 6			
XX	AAE24450			
ID	AAE24450 standard; protein; 1427 AA.			
AC	AAE24450;			
XX	04-OCT-2002 (first entry)			
DT	Human Von Willebrand factor-cleaving protease (vWF-cp).			
XX	Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;			
KW	transgenic animal; immunisation; thromboembolic disease; preclampsia;			
KW	thrombotic thrombocytic purpura; TTP; Henoch-Schonlein purpura;			
KW	thrombosis; neonatal thrombocytopaenia; haemolytic-uraemic syndrome;			
KW	transgenic; anticoagulant; chromosome 9.			
XX	Homo sapiens.			
OS	Key	Location/Qualifiers		
EH	Peptide	1..51		
FT	Protein	/label= Signal_peptide		
FT	Cleavage-site	52..1427		
FT	Region	/note= "Mature human vWF-cp protein"		
FT	Region	69..75		
FT	Region	/note= "Furin cleavage site"		
FT	Region	224..228		
FT	Region	/note= "Catalytical side"		
FT	Domain	249		
FT	Domain	/note= "Met turn"		
FT	Domain	301..377		
FT	Domain	/note= "Distintegrin like motif"		
FT	Region	387..439		
FT	Region	/note= "Thrombospondin type I motif"		
FT	Region	441..553		
FT	Region	/note= "Cysteine rich region"		
FT	Domain	554..687		
FT	Domain	/note= "Spacer"		
FT	Domain	688..743		
FT	Domain	/note= "Thrombospondin type I motif"		
FT	Domain	744..805		
FT	Domain	/note= "Thrombospondin type I motif"		
FT	Domain	897..952		
FT	Domain	/note= "Thrombospondin type I motif"		
FT	Domain	953..1013		
FT	Domain	/note= "Thrombospondin type I motif"		
FT	Domain	1016..1073		
FT	Domain	/note= "Thrombospondin type I motif"		
FT	Domain	1075..1131		
FT	Domain	/note= "Thrombospondin type I motif"		
XX	WO200242441-A2.			
XX	30-MAY-2002.			
XX	20-NOV-2001; 2001WO-EP013391.			
XX	22-NOV-2000; 2000US-00721254.			
XX	12-APR-2001; 2001US-00833328.			
XX	(BAXT) BAXTER AG.			
XX	Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;			
XX	Scheiflinger F, Antcaine G, Kerschbaumer R, Tagliavacca L;			
XX	Zimmermann K, Voelkel D;			

```
XX WPI; 2002-479950/51.
DR N-PSDB; AAD39332.
XX
XX Novel isolated or substantially purified Von Willebrand factor-cleaving
PT protease, useful for producing preparation for therapy of thrombosis and
PT thromboembolic disease such as thrombotic thrombocytopenic purpura.
XX
XX Claim 1; Fig 5; 93pp; English.
XX
XX The invention relates to an isolated or substantially pure Von Willebrand
CC factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
CC purifying vWF which involves providing vWF-cp as a ligand, contacting a
CC solution comprising vWF with the polypeptide ligand under conditions
CC where vWF is bound to the ligand and recovering from the ligand purified
CC vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
CC which involves immunising an animal with vWF-cp and isolating the anti-
CC vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
CC producing a preparation of prophylaxis and therapy of thrombosis and
CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
CC HemoCh-Schönlein purpura, pre-eclampsia, neonatal thrombocytopenia or
CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing
CC plasmatc or recombinantly produced vWF. The invention is useful for
CC construction expression systems and generating transgenic animals which
CC express the polypeptide in vivo. The present sequence is human vWF-cp
CC protein. vWF-cp gene is located on chromosome 9
XX
XX Sequence 1427 AA;
SQ
Query Match 100.0%; Score 1141; DB 5; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.2e-109;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLELVAVGPDVFOAQHEDTERVLTNINIGAEILLRDPISLGAQRFVHLVGMVIL 60
DB 75 AAGGILHLELVAVGPDVFOAQHEDTERVLTNINIGAEILLRDPISLGAQRFVHLVGMVIL 134
QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDDTPGHADLVLYITRFDELPDGNRQVRG 120
DB 135 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDDTPGHADLVLYITRFDELPDGNRQVRG 194
QY 121 VTQLGACSPWTSCLITETGFDLGVTTIAHEIGHSGFLEHDPAGSGCGPSGHVNASDGA 180
DB 195 VTQLGACSPWTSCLITETGFDLGVTTIAHEIGHSGFLEHDPAGSGCGPSGHVNASDGA 254
QY 181 APRAGLAWSPCSRRLSLLSAGRCVWDPPRP 214
DB 255 APRAGLAWSPCSRRLSLLSAGRCVWDPPRP 288
RESULT 7
ADD94038
ID ADD94038 standard; protein; 1427 AA.
XX
XX ADD94038;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human aggrecanase protein amino acid sequence.
XX
XX aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
XX cartilage degradation; osteoarthritis; inflammatory joint disease;
XX antiarthritic; osteopathic; antiinflammatory;
XX aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
XX human; enzyme.
XX
XX Homo sapiens.
XX
XX US2003105313-A1.
XX
XX 05-JUN-2003.
XX
XX 25-JAN-2002; 2002US-00057487.
```

```
XX 16-OCT-2001; 2001US-00978979.
XX
XX (AMHP ) AMERICAN HOME PROD CORP.
XX
XX Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
XX
XX WPI; 2003-801251/75.
XX
XX N-PSDB; ADD94037.
XX
XX New isolated DNA molecule encoding an aggrecanase polypeptide for
PT producing a purified human aggrecanase protein which can be used to
PT develop inhibitors of aggrecanase.
XX
XX Claim 14; SEQ ID NO 8; 24pp; English.
XX
XX This invention relates to a novel human aggrecanase protein and the DNA
CC sequence which encodes it. Aggrecan is a major extracellular component of
CC articular cartilage. It is a proteoglycan responsible for providing
CC cartilage with its mechanical properties of compressability and
CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
CC cleavage of aggrecan thereby having a role in cartilage degradation
CC associated with osteoarthritis and inflammatory joint disease. Compounds
CC which inhibit the activity of the protein of the invention may have
CC antiarthritic, osteopathic or antiinflammatory activity. The invention
CC may be used to produce a purified human aggrecanase protein. The protein
CC (or fragment) may be used to develop inhibitors of aggrecanase, using
CC three dimensional structural analysis or computer aided drug design. A
CC peptide which binds to aggrecanase is used to inhibit the proteolytic
CC degradation of aggrecan. The invention may be useful for the development
CC of therapeutics for the treatment of aggrecanase-associated disorders,
CC such as, osteoarthritis and other inflammatory conditions. The present
CC sequence is that of the human aggrecanase protein (full length sequence)
CC of the invention.
XX
XX Sequence 1427 AA;
SQ
Query Match 100.0%; Score 1141; DB 7; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.2e-109;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLELVAVGPDVFOAQHEDTERVLTNINIGAEILLRDPISLGAQRFVHLVGMVIL 60
DB 75 AAGGILHLELVAVGPDVFOAQHEDTERVLTNINIGAEILLRDPISLGAQRFVHLVGMVIL 134
QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDDTPGHADLVLYITRFDELPDGNRQVRG 120
DB 135 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDDTPGHADLVLYITRFDELPDGNRQVRG 194
QY 121 VTQLGACSPWTSCLITETGFDLGVTTIAHEIGHSGFLEHDPAGSGCGPSGHVNASDGA 180
DB 195 VTQLGACSPWTSCLITETGFDLGVTTIAHEIGHSGFLEHDPAGSGCGPSGHVNASDGA 254
QY 181 APRAGLAWSPCSRRLSLLSAGRCVWDPPRP 214
DB 255 APRAGLAWSPCSRRLSLLSAGRCVWDPPRP 288
RESULT 8
ABB98125
ID ABB98125 standard; protein; 1445 AA.
XX
XX ABB98125;
XX
XX 17-OCT-2002 (first entry)
XX
XX Human PMWV Incyte ID 7473607CD1.
XX
XX Human; PMWV; protein modification and maintenance molecule;
XX anticonvulsant; neuroprotective; nootropic; cyostatic; antipsoriatic;
XX antiaesthetic; dermatological; antidiabetic; antiparkinsonian;
XX antianaemic; antiinflammatory; antiulcer; antianginal; cardiant;
XX hepatotropic; osteopathic; antiemetic; antipyretic; virucide;
```


KW	antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
KW	haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
KW	cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
KW	anti-tumour; anti-rheumatic; immunosuppressive; anti-allergic; antithyroid;
KW	nephrotropic; anti-gout; thyromimetic; antiarthritic; uropachic;
KW	ophthalmological; anti-parasitic; tranquiliser; vulneary; keratolytic;
KW	auditory; antiseborrheic; antidepressant; neuroleptic; antinfertility;
KW	anthelmintic; protozoacide; Crohn's disease; hypertension; autoimmune;
KW	inflammatory; anaemia; cell proliferative; developmental; epithelial;
KW	scabies; neurological; Alzheimer's disease; reproductive;
KW	ectopic pregnancy; gene therapy; vaccine; disorder;
KW	procollagen I N-proteinase.
XX	
XX	Homo sapiens.
OS	
XX	WO200246383-A2.
PN	
XX	13-JUN-2002.
PD	
XX	
XX	05-DEC-2001; 2001WO-US04964.
PF	
XX	
XX	08-DEC-2000; 2000US-0254399P.
PR	
XX	21-DEC-2000; 2000US-0257803P.
PR	
XX	05-JAN-2001; 2001US-0260110P.
PR	
XX	19-JAN-2001; 2001US-0262851P.
PR	
XX	25-JAN-2001; 2001US-0264623P.
PR	
XX	(INCY-) INCYTE GENOMICS INC.
PA	
XX	
XX	Yue H, Azimzai Y, Kalliock DA, Baughn MR, Griffin JA, Swarnakar A;
PPI	Lal PG, Wallia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;
PPI	Ramkumar J, Thangavelu K, Lu Y, Warren BA, Lu DAM, Lee EA;
PPI	Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;
PPI	Sanjanwalla MW;
XX	
XX	WPI; 2002-519664/55.
DR	N-PSDB; ABQ75946.
DR	
XX	
XX	New isolated Protein Modification and Maintenance peptides, useful
PT	for diagnosis, and treatment of e.g. gastrointestinal disorders.
PT	
XX	
XX	Claim 1 (a); Page 154-157; 200pp; English.
PS	
XX	
XX	The invention relates to an isolated Protein Modification and Maintenance
CC	(PWM) polypeptide. Polypeptides of the invention may be used in the
CC	diagnosis, treatment and prevention of disorders associated with
CC	decreased expression or activity of PWM. These include gastrointestinal
CC	disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
CC	hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
CC	proliferative disorders, developmental disorders, epithelial disorders
CC	(e.g. scabies), neurological disorders (e.g. Alzheimer's disease)
CC	reproductive disorders (e.g. ectopic pregnancy), and in gene therapy or a
CC	vaccine for such diseases. They may also be used in the assessment of the
CC	effects of exogenous compound on the expression of nucleic acid and amino
CC	acid sequences of protein modification and maintenance molecules. The
CC	current sequence represents a human PWM of the invention, which has been
CC	found to have homology with B. taurus procollagen I N-proteinase
CC	
XX	Sequence 1445 AA;
SQ	
Query Match	100.0%; Score 1141; DB 5; Length 1445;
Best Local Similarity	100.0%; Pred. No. 1.2e-109;
Matches 214; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 AAGGILHLELVAVGPDVFQAHQEDTERVYLTNLNIGAEILLRDPSLGAQFRVHLVKWVIL 60
Db	78 AAGGILHLELVAVGPDVFQAHQEDTERVYLTNLNIGAEILLRDPSLGAQFRVHLVKWVIL 137
Qy	61 TEPEGAPNITANLTSSLVSVCWSQTINPEDDTPDGHADLVLYITRFDLELDPGNQVRG 120
Db	138 TEPEGAPNITANLTSSLVSVCWSQTINPEDDTPDGHADLVLYITRFDLELDPGNQVRG 197
Qy	121 VTQLGGASPTWSCILTEDTGFDLGVTIAEIGHSGFLEHDPAGPGSGCGPSGHVWASDGA 180

Db	198	VTQGGACSTNSCLITDTCFDLGVTTIAHEIGHSGFLEHDAFGSCGSPGHVMSDGA	257
Qy	181	APRAGLAWPCSRRLQLSLISAGRARCVMDPPRP	214
Db	258	APRAGLAWPCSRRLQLSLISAGRARCVMDPPRP	291
RESULT 9			
AAO16621			
ID	AAO16621	standard; protein; 312 AA.	
XX	AAO16621;		
XX	15-MAY-2003	(first entry)	
XX	Human von Willebrand factor (vWF)-cleaving enzyme-related protein #8.		
XX	Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;		
KW	von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;		
KW	myocardial infarction; cerebral infarction; arteriosclerosis;		
KW	platelet thrombosis; stenosis.		
XX	Homo sapiens.		
OS			
XX	WO200289366-A1.		
XX	07-NOV-2002.		
XX	25-APR-2002; 2002WO-JP004141.		
XX	25-APR-2001; 2001JP-00128342.		
PR	27-JUL-2001; 2001JP-00227510.		
PR	28-SEP-2001; 2001JP-00302977.		
PR	25-JAN-2002; 2002JP-00017596.		
XX	(KAGA) CHEMO-SERO-THERAPEUTIC RES INST.		
XX	Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;		
FI	WPI; 2003-120479/11.		
XX	N-PSDB; ABT32587.		
DR			
DR	von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and		
PT	supplementary therapy for, thrombotic thrombocytopenic purpura, and for		
PT	developing drugs for e.g. myocardial infarction and cerebral infarction.		
XX	Claim 4; Page 121-123; 144pp; Japanese.		
PS			
XX	The invention comprises the amino acid and coding sequence of a von		
CC	Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of		
CC	the invention are useful in the diagnosis and treatment of		
CC	thrombocytopenic purpura, and in developing drugs for myocardial		
CC	infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,		
CC	and stenosis. The present amino acid sequence represents a human von		
CC	Willebrand factor (vWF)-cleaving enzyme-related protein		
XX			
SQ	Sequence 312 AA;		
Query Match	93.3%;	Score 1065;	DB 6; Length 312;
Best Local Similarity	98.1%;	Pred. No. 1.3e-102;	
Matches 202;	Conservative 0;	Mismatches 4;	Indels 0; Gaps 0;
Qy	1	AAGGTHLELLVAVGPDVFQAHQEDTERVLTNLNIGALLRDPISLGAQFRVHLVQWVIL	60
Db	1	AAGGTHLELLVAVGPDVFQAHQEDTERVLTNLNIGALLRDPISLGAQFRVHLVQWVIL	60
Qy	61	TEPEGAPNITANLTSSLVCGWSQTINPEDTDPGHADLVLYITRFDLELDPGNQVRG	120
Db	61	TEPEGAPNITANLTSSLVCGWSQTINPEDTDPGHADLVLYITRFDLELDPGNQVRG	120
Qy	121	VTQLGACSPVWSCLITDTCFDLGVTTIAHEIGHSGFLEHDAFGSCGSPGHVMSDGA	180

Db 121 VTQLGACSTWSCLTEDTGFGLVGTIAHIGHSFGLHDPGSGCGPSGHVNASDGA 180

QY 181 APRAGLAWSPCSRRLSLLSAGRAR 206

Db 181 APRAGLAWSPCSRRLSLLRTGALR 206

RESULT 10

AAO16620

ID AAO16620 standard; protein; 1323 AA.

XX

AC AAO16620;

XX

DT 15-MAY-2003 (first entry)

XX

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #7.

XX

KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

KW platelet thrombosis; stenosis.

XX

OS Homo sapiens.

XX

PN WO200288366-A1.

XX

PD 07-NOV-2002.

XX

PF 25-APR-2002; 2002WO-JP004141.

XX

PR 25-APR-2001; 2001JP-00128342.

PR 27-JUL-2001; 2001JP-00227510.

PR 28-SEP-2001; 2001JP-00302977.

PR 25-JAN-2002; 2002JP-00017596.

XX

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX

PI Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX

DR WPI; 2003-120479/11.

DR N-PSDB; ABT32586.

XX

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and supplementary therapy for, thrombotic thrombocytopenic purpura, and for developing drugs for e.g. myocardial infarction and cerebral infarction.

PT

PS Claim 4; Page 112-121; 144pp; Japanese.

XX

CC The invention comprises the amino acid and coding sequence of a von Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of the invention are useful in the diagnosis and treatment of thrombocytopenic purpura, and in developing drugs for myocardial infarction, cerebral infarction, arteriosclerosis, platelet thrombosis, and stenosis. The present amino acid sequence represents a human von Willebrand factor (vWF)-cleaving enzyme-related protein

XX

SQ Sequence 1323 AA;

Query Match 93.3%; Score 1065; DB 6; Length 1323;

Best Local Similarity 98.1%; Pred. No. 9.9e-102;

Matches 202; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPISLGAQFRVHLVKVIL 60

Db 1 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPISLGAQFRVHLVKVIL 60

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELPDGNQVRG 120

Db 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELPDGNQVRG 120

QY 121 VTQLGACSTWSCLTEDTGFGLVGTIAHIGHSFGLHDPGSGCGPSGHVNASDGA 180

Db 121 VTQLGACSTWSCLTEDTGFGLVGTIAHIGHSFGLHDPGSGCGPSGHVNASDGA 180

QY 181 APRAGLAWSPCSRRLSLLSAGRAR 206

Db 181 APRAGLAWSPCSRRLSLLRTGALR 206

RESULT 11

AAO16619

ID AAO16619 standard; protein; 1378 AA.

XX

AC AAO16619;

XX

DT 15-MAY-2003 (first entry)

XX

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #6.

XX

KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

KW platelet thrombosis; stenosis.

XX

OS Homo sapiens.

XX

PN WO200288366-A1.

XX

PD 07-NOV-2002.

XX

PF 25-APR-2002; 2002WO-JP004141.

XX

PR 25-APR-2001; 2001JP-00128342.

PR 27-JUL-2001; 2001JP-00227510.

PR 28-SEP-2001; 2001JP-00302977.

PR 25-JAN-2002; 2002JP-00017596.

XX

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX

PI Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX

DR WPI; 2003-120479/11.

DR N-PSDB; ABT32585.

XX

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and supplementary therapy for, thrombotic thrombocytopenic purpura, and for developing drugs for e.g. myocardial infarction and cerebral infarction.

PT

PS Claim 4; Page 102-111; 144pp; Japanese.

XX

CC The invention comprises the amino acid and coding sequence of a von Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of the invention are useful in the diagnosis and treatment of thrombocytopenic purpura, and in developing drugs for myocardial infarction, cerebral infarction, arteriosclerosis, platelet thrombosis, and stenosis. The present amino acid sequence represents a human von Willebrand factor (vWF)-cleaving enzyme-related protein

XX

SQ Sequence 1378 AA;

Query Match 93.3%; Score 1065; DB 6; Length 1378;

Best Local Similarity 98.1%; Pred. No. 1e-101;

Matches 202; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPISLGAQFRVHLVKVIL 60

Db 1 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNLNIGAEILLRDPISLGAQFRVHLVKVIL 60

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELPDGNQVRG 120

Db 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELPDGNQVRG 120

QY 121 VTQLGACSTWSCLTEDTGFGLVGTIAHIGHSFGLHDPGSGCGPSGHVNASDGA 180

Db 121 VTQLGACSTWSCLTEDTGFGLVGTIAHIGHSFGLHDPGSGCGPSGHVNASDGA 180

QY 181 APRAGLAWSPCRRQLLSLSAGRAR 206
DB 181 APRAGLAWSPCRRQLLSLRTGALR 206

RESULT 12

AAO16622
ID AAO16622 standard; protein; 270 AA.

AC AAO16622;
XX

DT 15-MAY-2003 (first entry)
XX

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #9.
XX

KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
XX
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
XX
KW myocardial infarction; cerebral infarction; arteriosclerosis;
XX
KW platelet thrombosis; stenosis.
XX

OS Homo sapiens.
XX

PN WO200288366-A1.
XX

XX
PD 07-NOV-2002.
XX

XX 25-APR-2002; 2002WO-JP004141.
XX

XX 25-APR-2001; 2001JP-00128342.
XX

PR 27-JUL-2001; 2001JP-00227510.
XX

PR 28-SEP-2001; 2001JP-00302977.
XX

PR 25-JAN-2002; 2002JP-00017596.
XX

XX (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.
PA

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
PI

XX
DR WPI; 2003-120479/11.
DR

XX N-PSDB; ABT32588.
XX

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
PT developing drugs for e.g. myocardial infarction and cerebral infarction.
XX

PS Claim 4; Page 124-126; 144pp; Japanese.
XX

XX The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein
XX

SQ Sequence 270 AA;

Query Match 92.9%; Score 1059.5; DB 6; Length 270;
Best Local Similarity 95.8%; Pred. No. 3.8e-102; Indels 7; Gaps 2;
Matches 205; Conservative 0; Mismatches 2;

QY 1 AAGGILHLELVAVGPDVFQAEQDTERVLTNLNIGALLRDPSPSLGAQFRVHLVXVIL 60

DB 1 AAGGILHLELVAVGPDVFQAEQDTERVLTNLNIGALLRDPSPSLGAQFRVHLVXVIL 60

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPDGHADLVLYITRFDLELPDGNQVRG 120

DB 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPDGHADLVLYITRFDLELPDGNQVRG 120

QY 121 VTQLGGACSPWWSCLITETGFDLGVITIAHTECHSFGLEHDAFGSGCGPSGHVMSDGA 180

DB 121 VTQLGGACSPWWSCLITETGFDLGVITIAHTECHSFGLEHDAFGSGCGPSGHVMSDGA 180

QY 181 APRAGLAWSPCRRQLLSLSAGRARCVDPPRP 214

DB 181 APRAGLAWSPCRRQLLSLL-----RPV--PPSP 207

RESULT 13

AAU97641
ID AAU97641 standard; protein; 242 AA.

AC AAU97641;
XX

DT 27-AUG-2002 (first entry)
XX

DE Human aggrecanase protein.
XX

KW Aggrecanase; osteopathic; antiarthritic; antiinflammatory; human;
XX
KW cartilage; osteoarthritis; inflammatory disease; enzyme.
XX

OS Homo sapiens.
XX

PN WO200233093-A2.
XX

XX 25-APR-2002.
PD

XX 17-OCT-2001; 2001WO-US032458.
PF

XX 18-OCT-2000; 2000US-0241469P.
PR

XX (GEMY) GENETICS INST INC.
PA

XX Racie LA, Twine NC, Agostino MJ, Wolfman NM, Morris EA;
PI

XX WPI; 2002-454602/48.
XX

XX N-PSDB; ABK52579.
XX

XX Novel purified aggrecanase polypeptide useful for developing inhibitors
PT and antibodies to the aggrecanase polypeptide, which are useful for
PT treating aggrecanase-associated condition such as osteoarthritis.
XX

XX Claim 11; Page 33-34; 41pp; English.

XX This invention relates to the cDNA and protein sequences of a novel human
CC aggrecanase polypeptide. The protein of the invention may be used to
CC inhibit the proteolytic activity of aggrecanase, or to inhibit the
CC aggrecanase-mediated cleavage of aggrecan in cartilage. The protein of
CC the invention is useful for developing inhibitors of aggrecanase protein.
CC The cDNA sequence encoding the aggrecanase protein of the invention is
CC useful for designing probes for obtaining DNA sequences encoding other
CC aggrecanase molecules. The cDNA sequence is also useful for detecting
CC mRNA encoding aggrecanase in a given cell population, and thus for
CC detecting or diagnosing genetic disorders involving the aggrecanase, or
CC disorders involving cellular, organ or tissue disorders in which
CC aggrecanase is irregularly transcribed or expressed. The DNA sequences
CC may also be useful for preparing vectors for gene therapy applications.
CC An inhibitor of the protein is useful in treating conditions
CC characterised by degradation of articular cartilage, by blocking the
CC enzyme's proteolytic activity. An aggrecanase protein inhibitor and a
CC method for inhibition of its activity are useful for treating various
CC aggrecanase-associated conditions including osteoarthritis and other
CC inflammatory diseases. The present sequence represents the human
CC aggrecanase protein of the invention
XX

SQ Sequence 242 AA;

Query Match 92.6%; Score 1057; DB 5; Length 242;
Best Local Similarity 100.0%; Pred. No. 6e-102; Indels 0; Gaps 0;
Matches 200; Conservative 0; Mismatches 0;

QY 1 AAGGILHLELVAVGPDVFQAEQDTERVLTNLNIGALLRDPSPSLGAQFRVHLVXVIL 60

DB 42 AAGGILHLELVAVGPDVFQAEQDTERVLTNLNIGALLRDPSPSLGAQFRVHLVXVIL 101

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPDGHADLVLYITRFDLELPDGNQVRG 120

Db 102 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 161
QY 121 VTQLGACSPWTSCLITDGTGDLGVTTIAHEIGHSGFLEHDPGSGCGPSGHVWASDGA 180
Db 162 VTQLGACSPWTSCLITDGTGDLGVTTIAHEIGHSGFLEHDPGSGCGPSGHVWASDGA 221
QY 181 APRAGLAWSPCSRRLSLL 200
Db 222 APRAGLAWSPCSRRLSLL 241
RESULT 14
ID ADD94031 standard; protein; 242 AA.
XX ADD94031;
XX 29-JAN-2004 (first entry)
XX Human aggrecanase protein partial amino acid sequence.
XX aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
KW cartilage degradation; osteoarthritis; inflammatory joint disease;
KW antiarthritic; osteopathic; antiinflammatory;
KW aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
KW human; enzyme.
XX
OS Homo sapiens.
XX US2003105313-A1.
XX 05-JUN-2003.
XX 25-JAN-2002; 2002US-00057487.
XX 16-OCT-2001; 2001US-00978979.
XX (AMHP) AMERICAN HOME PROD CORP.
XX
XX Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
XX WPI; 2003-801251/75.
XX N-PSDB; ADD94032, ADD94033.
XX
XX New isolated DNA molecule encoding an aggrecanase polypeptide for
PT producing a purified human aggrecanase protein which can be used to
PT develop inhibitors of aggrecanase.
XX
XX Claim 13; SEQ ID NO 1; 24pp; English.
XX
XX This invention relates to a novel human aggrecanase protein and the DNA
CC sequence which encodes it. Aggrecan is a major extracellular component of
CC articular cartilage. It is a proteoglycan responsible for providing
CC cartilage with its mechanical properties of compressability and
CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
CC cleavage of aggrecan thereby having a role in cartilage degradation
CC associated with osteoarthritis and inflammatory joint disease. Compounds
CC which inhibit the activity of the protein of the invention may have
CC antiarthritic, osteopathic or antiinflammatory activity. The invention
CC may be used to produce a purified human aggrecanase protein. The protein
CC (or fragment) may be used to develop inhibitors of aggrecanase, using
CC three dimensional structural analysis or computer aided drug design. A
CC peptide which binds to aggrecanase is used to inhibit the proteolytic
CC degradation of aggrecan. The invention may be useful for the development
CC of therapeutics for the treatment of aggrecanase-associated disorders,
CC such as, osteoarthritis and other inflammatory conditions. The present
CC sequence is that of the human aggrecanase protein (partial) of the
CC invention.
XX
XX Sequence 242 AA;
XX
XX Query Match 92.6%; Score 1057; DB 7; Length 242;
XX Best Local Similarity 100.0%; Pred. No. 6e-102;

Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLLVAVGDFVQAHQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKMVL 60
Db 42 AAGGILHLLVAVGDFVQAHQEDTERYVLTNLNIGALLRDPGLGQAFRVHLVKMVL 101
QY 61 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 120
Db 102 TEPEGAPNITANTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 161
QY 121 VTQLGACSPWTSCLITDGTGDLGVTTIAHEIGHSGFLEHDPGSGCGPSGHVWASDGA 180
Db 162 VTQLGACSPWTSCLITDGTGDLGVTTIAHEIGHSGFLEHDPGSGCGPSGHVWASDGA 221
QY 181 APRAGLAWSPCSRRLSLL 200
Db 222 APRAGLAWSPCSRRLSLL 241
RESULT 15
AAG63826
ID AAG63826 standard; protein; 203 AA.
XX AAG63826;
XX 29-OCT-2001 (first entry)
DT
XX
DE Amino acid sequence of a human zdint5 polypeptide.
XX
XX Human; zdint5; anti-angiogenic; intestinal polypeptide; wound healing;
KW extracellular matrix interaction; tumour suppression; gamete maturation;
KW immunologic recognition; gastrointestinal irradiation; chemotherapy;
KW proteolysis; apoptosis; angiogenesis; infection; cell adhesion; trauma;
KW cell fusion; cell signalling; tumour; Crohn's disease; melanoma; trauma;
KW inflammatory bowel disease; food poisoning; degenerative disease;
KW inflammation; fertility; gamete maturation; epithelial disorder.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Modified-site 69
FT /note= "potential N-linked glycosylation site"
FT Modified-site 73
FT /note= "potential N-linked glycosylation site"
FT Region 151..161
FT /note= "zinc-binding motif"
XX
XX WO200159112-A1.
XX
XX 16-AUG-2001.
XX
XX 09-FEB-2001; 2001WO-US004198.
XX
XX 10-FEB-2000; 2000US-00501806.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX Holloway JL, Sheppard PO;
XX
XX WPI; 2001-522477/57.
XX N-PSDB; AAH74759.
XX
XX New anti-angiogenic intestinal polypeptides, zdint5 polypeptides, which
PT are members of disintegrin proteases, for modulating extracellular matrix
PT interaction, tumor suppression and wound healing.
XX
XX Claim 1; Page 73-74; 92pp; English.
XX
XX The present sequence represents a human zdint5 polypeptide. The zdint5
CC polypeptide is an anti-angiogenic intestinal polypeptide. Zdint5 is used
CC for modulating extracellular matrix interactions. Zdint5 polypeptide is
CC useful as a tool for identifying new family members of polypeptides.
CC Zdint5 polynucleotides are useful as probes or primers to clone 5' non-

CC coding regions of zdint5 gene. Zdint5 polypeptides are used for tumour
CC suppression, gamete maturation, immunologic recognition, and growth and
CC differentiation either working in isolation or in conjunction with other
CC molecules in colon, small intestine, fetal lung, testis and B-cells.
CC zdint5 polypeptides are also useful for promoting wound healing, in the
CC treatment of disorders associated with recovery after gastrointestinal
CC irradiation, chemotherapy or antibody use, as anti-infectives, and
CC extracellular matrix repair and remodeling. The polypeptides are also
CC useful for modulating proteolysis, apoptosis, angiogenesis, infection,
CC cell adhesion, cell fusion and signalling. The polypeptides are also
CC useful for treating tumour formation, Crohn's disease, inflammatory bowel
CC disease, food poisoning, melanoma, degenerative diseases, disorders
CC related to immunity, inflammation, fertility, gamete maturation,
CC immunology, trauma and epithelial disorders
XX
SQ Sequence 203 AA;

Query Match 91.7%; Score 1046; DB 4; Length 203;
Best Local Similarity 99.5%; Pred. No. 6.6e-101;
Matches 197; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNINIGALLRDPGLGAQFRVHLVKWVIL 60
DB 2 AAGGILHLELLVAVGPDVFOAQHEDTERYVLTNINIGALLRDPGLGAQFRVHLVKWVIL 61
QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 120
DB 62 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNRQVRG 121
QY 121 VTQLGGACSTWSCLITETGFDLGVTIAHIGHSFGLHGDGAPGCGPSCGHVNASDGA 180
DB 122 VTQLGGACSTWSCLITETGFDLGVTIAHIGHSFGLHGDGAPGCGPSCGHVNASDGA 181
QY 181 APRAGLAWSPCSRQLLS 198
DB 182 APRAGLAWSPCSRQLLS 199

Search completed: March 13, 2004, 07:39:01
Job time : 32.3198 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 16.9417 Seconds
(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311
Perfect score: 1141
Sequence: 1 AAGGILHLELLVAVGPDVFO.....QLLSLLSAGRARCWVDPFRP 214

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

ALIGNMENTS

RESULT 1
US-10-222-334-4
; Sequence 4, Application US/10222334
; Publication NO: US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
; FILE REFERENCE: UM-07288
; CURRENT APPLICATION NUMBER: US/10/222.334
; PRIOR FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-4

Query Match		100.0%;	Score 1141;	DB 14;	Length 842;
Best Local Similarity		100.0%;	Pred. No. 2,7e-108;		
Matches 214;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AAGGILHLELLVAVGPDVFOAHOEDTERTYVLTNINIGAEELRDPISLGAQFRVHLVQWVIL	60		
Db	75	AAGGILHLELLVAVGPDVFOAHOEDTERTYVLTNINIGAEELRDPISLGAQFRVHLVQWVIL	134		
QY	61	TEPEGAPNITANITSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNQVRG	120		
Db	135	TEPEGAPNITANITSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDELDPGNQVRG	194		
QY	121	VTOLGGACSTWSCLITEDTGFDLGVITAEIGHSGFGLHDPGPGCGPSGHVWASDGA	180		
Db	195	VTOLGGACSTWSCLITEDTGFDLGVITAEIGHSGFGLHDPGPGCGPSGHVWASDGA	254		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1141	100.0	842	14	US-10-222-334-4
2	1141	100.0	1416	9	US-09-836-712-2
3	1141	100.0	1427	14	US-10-222-334-2
4	1141	100.0	1427	14	US-10-057-487-6
5	1057	92.6	242	9	US-09-978-979-1
6	1057	92.6	242	14	US-10-057-487-1
7	1046	91.7	203	9	US-09-781-080B-2
8	1046	91.7	1120	9	US-09-781-080B-11
9	773	67.7	148	9	US-09-833-328-4
10	773	67.7	150	9	US-09-833-328-6
11	716	62.8	136	9	US-09-833-328-15
12	702	61.5	133	9	US-09-833-328-2
13	361	31.6	997	11	US-09-981-151A-32
14	361	31.6	1686	15	US-10-386-414-2
15	360	31.6	997	9	US-09-918-171A-7

```
QY 181 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 214
DB 255 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 288

RESULT 2
US-09-836-712-2
; Sequence 2, Application US/09835712
; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1416
; TYPE: PRT
; ORGANISM: Human
US-09-836-712-2

Query Match 100.0%; Score 1141; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 5.4e-108;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 60
DB 98 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 157

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 120
DB 158 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 217

QY 121 VTQLGACSPWTSCLITETGFDLGVTIAHIGHSFGLHEDGAPGSGCGPSGHVWASDGA 180
DB 218 VTQLGACSPWTSCLITETGFDLGVTIAHIGHSFGLHEDGAPGSGCGPSGHVWASDGA 277

QY 181 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 214
DB 278 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 311

RESULT 3
US-10-222-334-2
; Sequence 2, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Galia
; TITLE OF INVENTION: Tsai, Han-Mou
; FILE REFERENCE: UM-07288
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2

Query Match 100.0%; Score 1141; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 5.5e-108;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 60

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 120
DB 75 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 134

QY 135 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 194
DB 181 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 214
DB 255 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 288

RESULT 4
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrecanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-057-487-8

Query Match 100.0%; Score 1141; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 5.5e-108;
Matches 214; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 60
DB 75 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGACQFRVHLVKWVIL 134

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 120
DB 135 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELDPGNGRQVRG 194

QY 121 VTQLGACSPWTSCLITETGFDLGVTIAHIGHSFGLHEDGAPGSGCGPSGHVWASDGA 180
DB 195 VTQLGACSPWTSCLITETGFDLGVTIAHIGHSFGLHEDGAPGSGCGPSGHVWASDGA 254

QY 181 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 214
DB 255 APRAGLAWPCSRRLQLLSLSAGRARCVDPPRP 288

RESULT 5
US-09-978-979-1
; Sequence 1, Application US/09978979
; Patent No. US20020151702A1
; GENERAL INFORMATION:
; APPLICANT: Racie, Lisa, A.
; Twine, Natalie, C.
; Agostino, Michael, J.
; Wolfman, Neil
; Morris, Elisabeth
; TITLE OF INVENTION: Aggrecanase Molecules
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; 
```

COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/978,979
FILING DATE: 16-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: GI 5435P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 242 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-978-979-1

Query Match	92.6%;	Score 1057;	DB 9;	Length 242;
Best Local Similarity	100.0%;	Pred. No. 2.1e-100;		
Matches 200;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
QY	1	AAGILHLELLVAGPDVFQAHQDTERYVLTNLNIGAEILRDSLGAFRVHLVKWVIL	60	
Db	42	AAGILHLELLVAGPDVFQAHQDTERYVLTNLNIGAEILRDSLGAFRVHLVKWVIL	101	
QY	61	TEPEGANITANTSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLELPDGNQVRG	120	
Db	102	TEPEGANITANTSSLLSCVGSQTINPEDDTPGHADLVLYITRFDLELPDGNQVRG	161	
QY	121	VTQGGACSPWCSCLITDGTDFDLGVTTAHEIGHSGFLEHDPAGSGCGPSGHVWASDGA	180	
Db	162	VTQGGACSPWCSCLITDGTDFDLGVTTAHEIGHSGFLEHDPAGSGCGPSGHVWASDGA	221	
QY	181	APRAGLAWSPCSRRLQLSLL	200	
Db	222	APRAGLAWSPCSRRLQLSLL	241	

```

RESULT 6
US-10-057-487-1
; Sequence 1, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggreacanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-057-487-1
Query Match          92.6%;  Score 1057;  DB 14;  Length 242;

```

[illegible]

```

RESULT 7
US-09-781-080B-2
; Sequence 2, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: YAMAMOTO, GAYLE
; TITLE OF INVENTION: Anti-Angiogenic Intestinal Peptides,
; TITLE OF INVENTION: zdint5
; FILE REFERENCE: 99-82
; CURRENT APPLICATION NUMBER: US/09/781,080B
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-781-080B-2

```

Query Match	91.7%	Score 1046;	DB 9;	Length 203;
Best Local Similarity	99.5%;	Pred. No. 2.3e-99;		
Matches 197;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAGGILHLELLVAGDPVFOAQHEDTERYVLTNLIGAEHLRDPSPSLGAQRFVHLVKMVL	60	
Db	2	AAGGILHLELLVAGDPVFOAQHEDTERYVLTNLIGAEHLRDPSPSLGAQRFVHLVKMVL	61	
Qy	61	TEPEGANITANTSSLLSVCGHSQTINPEDDTPDGHADLVLTITRFDLELPDGNRVRG	120	
Db	62	TEPOGAPNITANTSSLLSVCGHSQTINPEDDTPDGHADLVLTITRFDLELPDGNRVRG	121	
Qy	121	VTQJGGACSPWNSCLITDGFDLGVTTIAEIGHSGFGLHDGAPGSCGSPGHVWASDGA	180	
Db	122	VTQJGGACSPWNSCLITDGFDLGVTTIAEIGHSGFGLHDGAPGSCGSPGHVWASDGA	181	
Qy	181	APRAGLAWSPCSRQLLS	198	
Db	182	APRAGLAWSPCSRQLLS	199	

```

RESULT 8
US-09-781-080B-11
; Sequence 11, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: YAMAMOTO, GAYLE
; TITLE OF INVENTION: Anti-Angiogenic Intestinal peptides
; TITLE OF INVENTION: zdhntc5
; FILE REFERENCE: 99-82

```


CURRENT APPLICATION NUMBER: US/09/781,080B
CURRENT FILING DATE: 2002-01-30
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 1120
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1120)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-781-080B-11

Query Match
Best Local Similarity 91.7%; Score 1046; DB 9; Length 1120;
Matches 197; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148
DB 105 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120
DB 165 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120

QY 121 VTQLGGACSPWTSCLITETGFDLGVTI 148
DB 121 VTQLGGACSPWTSCLITETGFDLGVTI 148

RESULT 10
US-09-833-328-6
; Sequence 6, Application US/09833328
; Patent No. US20020136713A1
; GENERAL INFORMATION:
; APPLICANT: Laemmle, Bernhard
; APPLICANT: Schwarz, Hans-Peter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Antoine, Gerhard
; APPLICANT: Kerschbaumer, Randolph
; APPLICANT: Tagliavacca, Luigina
; APPLICANT: Zimmermann, Klaus
; APPLICANT: Furlan, Miha
; APPLICANT: Turecek, Peter
; APPLICANT: Gerritsen, Helena E.
; TITLE OF INVENTION: Composition Exhibiting a von Willebrand Factor (vWF) Protease A
; TITLE OF INVENTION: Comprising a Polypeptide Chain with the Amino Acid Sequence AA
; FILE REFERENCE: 247.00CIP
; CURRENT APPLICATION NUMBER: US/09/833,328
; PRIOR FILING DATE: 2001-04-12
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 6
; LENGTH: 150
; TYPE: PRT
; ORGANISM: human
US-09-833-328-6

Query Match
Best Local Similarity 67.7%; Score 773; DB 9; Length 150;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148
DB 3 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120
DB 63 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120

QY 121 VTQLGGACSPWTSCLITETGFDLGVTI 148
DB 123 VTQLGGACSPWTSCLITETGFDLGVTI 148

RESULT 11
US-09-833-328-15
; Sequence 15, Application US/09833328
; Patent No. US20020136713A1
; GENERAL INFORMATION:
; APPLICANT: Laemmle, Bernhard
; APPLICANT: Schwarz, Hans-Peter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Antoine, Gerhard
; APPLICANT: Kerschbaumer, Randolph
; APPLICANT: Tagliavacca, Luigina
; APPLICANT: Zimmermann, Klaus
; APPLICANT: Furlan, Miha
; APPLICANT: Turecek, Peter
; APPLICANT: Gerritsen, Helena E.
; TITLE OF INVENTION: Composition Exhibiting a von Willebrand Factor (vWF) Protease A
; TITLE OF INVENTION: Comprising a Polypeptide Chain with the Amino Acid Sequence AA

CURRENT APPLICATION NUMBER: US/09/781,080B
CURRENT FILING DATE: 2002-01-30
NUMBER OF SEQ ID NOS: 12
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 11
LENGTH: 1120
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)...(1120)
OTHER INFORMATION: Xaa = Any Amino Acid
US-09-781-080B-11

Query Match
Best Local Similarity 91.7%; Score 1046; DB 9; Length 1120;
Matches 197; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148
DB 105 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148

QY 61 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120
DB 165 TEPEGAPNITANLTSSLLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGDGNQVRG 120

QY 121 VTQLGGACSPWTSCLITETGFDLGVTI 148
DB 121 VTQLGGACSPWTSCLITETGFDLGVTI 148

RESULT 9
US-09-833-328-4
; Sequence 4, Application US/09833328
; Patent No. US20020136713A1
; GENERAL INFORMATION:
; APPLICANT: Laemmle, Bernhard
; APPLICANT: Schwarz, Hans-Peter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Antoine, Gerhard
; APPLICANT: Kerschbaumer, Randolph
; APPLICANT: Tagliavacca, Luigina
; APPLICANT: Zimmermann, Klaus
; APPLICANT: Furlan, Miha
; APPLICANT: Turecek, Peter
; APPLICANT: Gerritsen, Helena E.
; TITLE OF INVENTION: Composition Exhibiting a von Willebrand Factor (vWF) Protease A
; TITLE OF INVENTION: Comprising a Polypeptide Chain with the Amino Acid Sequence AA
; FILE REFERENCE: 247.00CIP
; CURRENT APPLICATION NUMBER: US/09/833,328
; PRIOR FILING DATE: 2001-04-12
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 148
; TYPE: PRT
; ORGANISM: human
US-09-833-328-4

Query Match
Best Local Similarity 67.7%; Score 773; DB 9; Length 148;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148
DB 1 AAGGILHLELLVAVGPDVFOAQHEDTERVYLTNLNIGAEILLRDPGLGVTI 148

```
; FILE REFERENCE: 247.00CIP
; CURRENT APPLICATION NUMBER: US/09/833,328
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 09/721,254
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 15
; LENGTH: 136
; TYPE: PRT
; ORGANISM: human
US-09-833-328-15

Query Match      62.8%; Score 716; DB 9; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.2e-65;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AVGPDVFOAQHEDTERVLTNINIGALLRDPSPSLGAQFRVHLVKXVILTEPEGAPNITAN 72
DB 1 AVGPDVFOAQHEDTERVLTNINIGALLRDPSPSLGAQFRVHLVKXVILTEPEGAPNITAN 60
QY 73 LTSSLLSVCGWSQTINPDDTDPGHADLVLYITRFDLELPDGNRQVRGVTQLGGACSTPW 132
DB 61 LTSSLLSVCGWSQTINPDDTDPGHADLVLYITRFDLELPDGNRQVRGVTQLGGACSTPW 120
QY 133 SCLITDGTGFDLGVTI 148
DB 121 SCLITDGTGFDLGVTI 136

RESULT 12
US-09-833-328-2
; Sequence 2, Application US/09833328
; Patent No. US20020136713A1
; GENERAL INFORMATION:
; APPLICANT: Laemmle, Bernhard
; APPLICANT: Schwarz, Hans-Peter
; APPLICANT: Scheiflinger, Friedrich
; APPLICANT: Antoine, Gerard
; APPLICANT: Kerschbaumer, Randolph
; APPLICANT: Tagliavacca, Luigina
; APPLICANT: Zimmermann, Klaus
; APPLICANT: Furlan, Miha
; APPLICANT: Turecek, Peter
; APPLICANT: Gerritsen, Helena E.
; TITLE OF INVENTION: Compositing a von Willebrand Factor (vWF) Protease Act
; TITLE OF INVENTION: Compositing a Polypeptide Chain with the Amino Acid Sequence AAGD
; FILE REFERENCE: 247.00CIP
; CURRENT APPLICATION NUMBER: US/09/833,328
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 09/721,254
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 133
; TYPE: PRT
; ORGANISM: human
US-09-833-328-2

Query Match      61.5%; Score 702; DB 9; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.1e-64;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 PDVFOAQHEDTERVLTNINIGALLRDPSPSLGAQFRVHLVKXVILTEPEGAPNITANLTS 75
DB 1 PDVFOAQHEDTERVLTNINIGALLRDPSPSLGAQFRVHLVKXVILTEPEGAPNITANLTS 60
QY 76 SLLSVCGWSQTINPDDTDPGHADLVLYITRFDLELPDGNRQVRGVTQLGGACSTPWSCL 135
DB 61 SLLSVCGWSQTINPDDTDPGHADLVLYITRFDLELPDGNRQVRGVTQLGGACSTPWSCL 120
QY 136 ITDGTGFDLGVTI 148
```

```
DB 121 ITDGTGFDLGVTI 133

RESULT 13
US-09-981-151A-32
; Sequence 32, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glennda
; APPLICANT: Miller, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shimkets, Richard A
; APPLICANT: Radigaru, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Kekuda, Rameesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-151A-32

Query Match      31.6%; Score 361; DB 11; Length 997;
Best Local Similarity 38.4%; Pred. No. 5.8e-28;
Matches 81; Conservative 35; Mismatches 87; Indels 8; Gaps 5;

QY 8 LELLVAVGPVFOAH-OEDTERVYLTNINIGALLRDPSPSLGAQFRVHLVKXVILTEPEGA 66
DB 244 VELLVADAKMVEYHGQFQVESYVLTNINMAGLFHDSIGNFIHITIVRLVULLEDEED 303
QY 67 PNITANLTSSLLSVCGWSQTINPDDTDPGHADLVLYITRFDLELPDGNR--QVRGVTQL 124
```

Db 304 LKITHADNTLKSFCWKQKSNKMGDAHPLHDDTAILLTRKDL-CAAMNRCETLGLSHV 362
QY 125 GGACPTWSCLITETDGLVGTIAHIGHSFGLHGDGAPSGCGPSG---HVMASDGA 181
Db 363 AGMCQPHRSCSINEDTGLPLAFTVAHELHSGFQIHDGS-GNDCEPVGKRPFIMSPOLLY 421
QY 182 PRAGLAWSPCSRQLLSLSAGRARCWDDPP 212
Db 422 DAAPLTWSRCRQYITRFLDRGWGLCLDDPP 452

RESULT 14
US-10-386-414-2
; Sequence 2, Application US/10386414
; Publication No. US2004006016A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Robison, Keith E.
; APPLICANT: White, David
; APPLICANT: Williamson, Mark W.
; APPLICANT: Cook, William James
; APPLICANT: Meyers, Rachel E.
; APPLICANT: MacBeth, Kyle J.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Chun, Miyoung
; TITLE OF INVENTION: NOVEL 27875, 22025, 27420, 17906, 16319,
; 55092 AND 10218 MOLECULES AND USES THEREFOR
; FILE REFERENCE: WPI03-0210MIM
; CURRENT APPLICATION NUMBER: US/10/386,414
; CURRENT FILING DATE: 2003-03-11
; PRIOR APPLICATION NUMBER: 09/426,282
; PRIOR FILING DATE: 1999-10-25
; PRIOR APPLICATION NUMBER: 09/668,266
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/330,970
; PRIOR FILING DATE: 1999-06-11
; PRIOR APPLICATION NUMBER: 09/724,599
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/860,193
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 10/283,023
; PRIOR FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: 60/335,044
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/010,943
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: 60/254,037
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 09/833,082
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1686
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-386-414-2

Query Match 31.6%; Score 361; DB 15; Length 1686;
Best Local Similarity 38.4%; Pred. No. 1.2e-27;
Matches 81; Conservative 35; Mismatches 87; Indels 8; Gaps 5;
QY 8 LELLVAVGPDVFQAH-QEDTERVYLTNLMNIGALLRDPGLGAGQFRVHLVKMVLTEPEGA 66
Db 244 VETLVADAKMVEYHGQPVESYVLTIMNVAGLFHPDPSIGNPIHITIVRLVLEDEED 303
QY 67 PNITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGNR--QVRGVTQL 124
Db 304 LKITHADNTLKSFCWKQKSNKMGDAHPLHDDTAILLTRKDL-CAAMNRCETLGLSHV 362
QY 125 GGACPTWSCLITETDGLVGTIAHIGHSFGLHGDGAPSGCGPSG---HVMASDGA 181

Db 363 AGMCQPHRSCSINEDTGLPLAFTVAHELHSGFQIHDGS-GNDCEPVGKRPFIMSPOLLY 421
QY 182 PRAGLAWSPCSRQLLSLSAGRARCWDDPP 212
Db 422 DAAPLTWSRCRQYITRFLDRGWGLCLDDPP 452
RESULT 15
US-09-918-171A-7
; Sequence 7, Application US/09918171A
; Patent No. US20020110894A1
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/04193
; CURRENT APPLICATION NUMBER: US/09/918,171A
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/369,364
; PRIOR FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-7
US-09-918-171A-7
Query Match 31.6%; Score 360; DB 9; Length 997;
Best Local Similarity 38.6%; Pred. No. 7.4e-28;
Matches 81; Conservative 34; Mismatches 87; Indels 8; Gaps 5;
QY 9 ELLVAVGPDVFQAH-QEDTERVYLTNLMNIGALLRDPGLGAGQFRVHLVKMVLTEPEGA 67
Db 245 ETLVADAKMVEYHGQPVESYVLTIMNVAGLFHPDPSIGNPIHITIVRLVLEDEEDL 304
QY 68 NITANLTSSLSVCGWSQTINPEDDTPGHADLVLYITRFDLELPGNR--QVRGVTQLG 125
Db 305 KITHADNTLKSFCWKQKSNKMGDAHPLHDDTAILLTRKDL-CAAMNRCETLGLSHVA 363
QY 126 GACPTWSCLITETDGLVGTIAHIGHSFGLHGDGAPSGCGPSG---HVMASDGAAP 182
Db 364 GMCPHRSINEDTGLPLAFTVAHELHSGFQIHDGS-GNDCEPVGKRPFIMSPOLLY 422
QY 183 RAGLAWSPCSRQLLSLSAGRARCWDDPP 212
Db 423 DAAPLTWSRCRQYITRFLDRGWGLCLDDPP 452

Search completed: March 13, 2004, 08:07:03
Job time : 17.9417 secs

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:34:04 ; Search time 9.02812 Seconds
(without alignments)
1223.727 Million cell updates/sec

Title: US-09-836-712-2_COPY_98_311
Perfect score: 1141
Sequence: 1 AAGGILHLLVAVGPDVFG.....QLLSLSAGRACVWDPPR 214

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	360	31.6	937	4	US-09-369-364A-7
2	325.5	28.5	356	4	US-09-963-791-20
3	325.5	28.5	438	4	US-09-963-791-22
4	325.5	28.5	507	4	US-09-963-791-10
5	325.5	28.5	589	4	US-09-963-791-12
6	325.5	28.5	757	4	US-09-963-791-24
7	325.5	28.5	908	4	US-09-963-791-2
8	307.5	27.0	1224	4	US-09-930-872-4
9	304.5	26.7	727	4	US-09-445-023A-1
10	304.5	26.7	949	4	US-09-568-559-2
11	304.5	26.7	967	4	US-09-130-491-2
12	301	26.4	317	4	US-09-963-791-16
13	301	26.4	468	4	US-09-963-791-6
14	299.5	26.2	727	4	US-09-445-023A-12
15	295	25.9	2165	4	US-09-800-729-155
16	288.5	25.3	1882	4	US-09-369-364A-13
17	284	24.9	837	4	US-09-122-126B-2
18	284	24.9	837	4	US-09-634-286A-2
19	273.5	24.0	491	4	US-09-930-872-2
20	272	23.8	1205	4	US-09-491-522-11
21	271	23.8	566	4	US-09-491-522-7
22	261	22.9	1211	4	US-09-491-522-5
23	256	22.4	905	4	US-09-369-364A-9
24	255.5	22.4	930	4	US-09-369-364A-2
25	255	22.3	874	4	US-09-369-364A-15
26	253	22.2	245	4	US-09-369-364A-11
27	252	22.1	608	4	US-09-130-491-13

28 249 21.8 859 4 US-09-369-364A-5 Sequence 5, Appli
29 245.5 21.5 930 4 US-09-122-126B-15 Sequence 15, Appl
30 245.5 21.5 930 4 US-09-634-286A-15 Sequence 15, Appl
31 233.5 20.5 1081 4 US-09-369-364A-17 Sequence 17, Appl
32 135.5 11.9 814 3 US-09-813-819-4 Sequence 4, Appli
33 135.5 11.9 814 4 US-09-920-048-4 Sequence 4, Appli
34 135.5 11.9 855 3 US-09-813-819-2 Sequence 2, Appli
35 135.5 11.9 855 4 US-09-920-048-2 Sequence 2, Appli
36 133.5 11.7 746 4 US-09-548-797B-4 Sequence 4, Appli
37 133.5 11.7 787 4 US-09-548-797B-5 Sequence 5, Appli
38 133.5 11.7 802 4 US-09-632-098-2 Sequence 2, Appli
39 133.5 11.7 812 4 US-09-632-098-4 Sequence 4, Appli
40 133.5 11.7 849 4 US-09-548-797B-6 Sequence 6, Appli
41 133 11.7 613 4 US-09-026-001A-10 Sequence 10, Appl
42 133 11.7 621 4 US-09-026-001A-18 Sequence 18, Appl
43 130 11.4 781 4 US-09-738-946-8 Sequence 8, Appli
44 120 10.5 201 3 US-09-411-329C-1 Sequence 1, Appli
45 120 10.5 201 4 US-09-411-335-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 997
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-7
US-09-369-364A-7

Query Match 31.6%; Score 360; DB 4; Length 997;
Best Local Similarity 38.6%; Pred. No. 2.6e-30;
Matches 81; Conservative 34; Mismatches 87; Indels 8; Gaps 5;
QY 9 ELLVAGPDVQAH-QEDTERYVLTNINIGALLRDPGLGAGQFRVHLVWVILTEPEGAP 67
245 ETLVADAKMVEYHGQPVESYVLTIMNVAAGLFHDPESIGNPIHITIVRLVLEDEEDL 304
QY 68 NITANLTSSLLSCGWSQTINPEDDTPGHADLVLYITRFDLELPDGNR--QVRGVYQLG 125
DB 305 KITHADNTLKSFKWQKSNMKGDAPHLHDDTALLTRKDL-CAAMNRPCTGLGSHVA 363
QY 126 GACSPWSCLTEDTGTDLGVTTAHEIGHGFLGHGAPGSGCGPSG---HYMASDGAAP 182
DB 364 GACQPRSSINEDTGLPLAFTVAHELGHGFGIHDGS-GNDCEPVGRKPFIMSPQLLYD 422
QY 183 RAGLAWSPCRRQLLSLSAGRACVWDPP 212
DB 423 AAPLTWSCRQYITRFLDRGWLGLDPP 452

RESULT 2
US-09-963-791-20
; Sequence 20, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John

APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 20
LENGTH: 356
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-20

Query Match 28.5%; Score 325.5; DB 4; Length 356;
Best Local Similarity 36.4%; Pred. No. 3.4e-27;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
DB 101 VETLVVADKMMVGYHGKDIHILSVNMIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLLSVCWSQTI---NPEDDTP-----GHADLVLYITRFDL-----ELPD 113
DB 159 PNLEINHADKSLDSFCWKQKSIHSHQSDGNTIPENGIAHHDNAVLITRYDICTYKNKPC 218
QY 114 GNRQVRGVTQLGACSPWCLITEGTGFDLGVTTIAHEIGHSGFLEHGDGAPSGCGPSGH 173
DB 219 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHEIGHFGNMHDDGI-GNSCGTKGH 274
QY 174 ----VNASDGAAPRAGLAWSPCSRRLSLLSAGRARCVCV-WDPPR 213
DB 275 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 319

RESULT 3
US-09-963-791-22
Sequence 22, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 438
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-22

Query Match 28.5%; Score 325.5; DB 4; Length 438;
Best Local Similarity 36.4%; Pred. No. 4.6e-27;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
DB 101 VETLVVADKMMVGYHGKDIHILSVNMIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLLSVCWSQTI---NPEDDTP-----GHADLVLYITRFDL-----ELPD 113
DB 159 PNLEINHADKSLDSFCWKQKSIHSHQSDGNTIPENGIAHHDNAVLITRYDICTYKNKPC 218

QY 114 GNRQVRGVTQLGACSPWCLITEGTGFDLGVTTIAHEIGHSGFLEHGDGAPSGCGPSGH 173
DB 219 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHEIGHFGNMHDDGI-GNSCGTKGH 274
QY 174 ----VNASDGAAPRAGLAWSPCSRRLSLLSAGRARCVCV-WDPPR 213
DB 275 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 319

RESULT 4
US-09-963-791-10
Sequence 10, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the S
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 507
TYPE: PRT
ORGANISM: Homo sapiens
US-09-963-791-10

Query Match 28.5%; Score 325.5; DB 4; Length 507;
Best Local Similarity 36.4%; Pred. No. 5.8e-27;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
DB 252 VETLVVADKMMVGYHGKDIHILSVNMIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 309
QY 67 PNITAN--LTSSLLSVCWSQTI---NPEDDTP-----GHADLVLYITRFDL-----ELPD 113
DB 310 PNLEINHADKSLDSFCWKQKSIHSHQSDGNTIPENGIAHHDNAVLITRYDICTYKNKPC 369
QY 114 GNRQVRGVTQLGACSPWCLITEGTGFDLGVTTIAHEIGHSGFLEHGDGAPSGCGPSGH 173
DB 370 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHEIGHFGNMHDDGI-GNSCGTKGH 425
QY 174 ----VNASDGAAPRAGLAWSPCSRRLSLLSAGRARCVCV-WDPPR 213
DB 426 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 470

RESULT 5
US-09-963-791-12
Sequence 12, Application US/09963791
Patent No. 6649399
GENERAL INFORMATION:
APPLICANT: Donoho, Gregory
APPLICANT: Turner, C. Alexander Jr.
APPLICANT: Friedrich, Glenn
APPLICANT: Scoville, John
APPLICANT: Zambrowicz, Brian
APPLICANT: Sands, Arthur T.
TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the S
FILE REFERENCE: LEX-0105-USA
CURRENT APPLICATION NUMBER: US/09/963,791
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/169,769
PRIOR FILING DATE: 1999-12-09

; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-12

Query Match 28.5%; Score 325.5; DB 4; Length 589;
Best Local Similarity 36.4%; Pred. No. 7.2e-27;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVYLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
Db 252 VETLVADKMMVGYHGKDIIEHYLSVMNIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 309
QY 67 PNITAN--LTSSLSVCGWSQTI---NPEDDTPD---GHADLVLYITRFDL----ELPD 113
Db 310 PNLNINHADKSLDSFCKWQKSLSHQSDGNTIPENGIAHDNAVLTIRYDICTYKMKPC 369
QY 114 GNRQVRGVTQLGACSPWTSCLITDGFGLGVTTIAHIGHSFGLHGDGAPSGCGPSGH 173
Db 370 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHIGHNFGMNHDI--GNSCGTKGH 425
QY 174 ----VNASDGAAPRAGLAWPCSRRLSLSAGRARCVCWDP 213
Db 426 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 470

RESULT 6
US-09-963-791-24
; Sequence 24, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 757
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-24

Query Match 28.5%; Score 325.5; DB 4; Length 757;
Best Local Similarity 36.4%; Pred. No. 1e-26;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVYLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
Db 101 VETLVADKMMVGYHGKDIIEHYLSVMNIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLSVCGWSQTI---NPEDDTPD---GHADLVLYITRFDL----ELPD 113
Db 159 PNLNINHADKSLDSFCKWQKSLSHQSDGNTIPENGIAHDNAVLTIRYDICTYKMKPC 218
QY 114 GNRQVRGVTQLGACSPWTSCLITDGFGLGVTTIAHIGHSFGLHGDGAPSGCGPSGH 173
Db 219 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHIGHNFGMNHDI--GNSCGTKGH 274
QY 174 ----VNASDGAAPRAGLAWPCSRRLSLSAGRARCVCWDP 213
Db 275 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 319

RESULT 7
US-09-963-791-2
; Sequence 2, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 908
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-2

Query Match 28.5%; Score 325.5; DB 4; Length 908;
Best Local Similarity 36.4%; Pred. No. 1.4e-26;
Matches 82; Conservative 36; Mismatches 82; Indels 25; Gaps 10;

QY 8 LELLVAVGPDVFOAH-QEDTERVYLTNLNIGAEILLRDPISLGAQFRVHLVKMVLTEPEGA 66
Db 252 VETLVADKMMVGYHGKDIIEHYLSVMNIVAKLYRDSLSLGNVNIIVARLIVLTEDQ-- 309
QY 67 PNITAN--LTSSLSVCGWSQTI---NPEDDTPD---GHADLVLYITRFDL----ELPD 113
Db 310 PNLNINHADKSLDSFCKWQKSLSHQSDGNTIPENGIAHDNAVLTIRYDICTYKMKPC 369
QY 114 GNRQVRGVTQLGACSPWTSCLITDGFGLGVTTIAHIGHSFGLHGDGAPSGCGPSGH 173
Db 370 G---TLGLASVAGMCEPERSCSINEDIGLSAFTIAHIGHNFGMNHDI--GNSCGTKGH 425
QY 174 ----VNASDGAAPRAGLAWPCSRRLSLSAGRARCVCWDP 213
Db 426 EAAKLMAAHITANTNPFWSACSRSYITSLDGRGTCLDNEPPK 470

RESULT 8
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Brin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; PRIOR FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match 27.0%; Score 307.5; DB 4; Length 1224;
Best Local Similarity 36.2%; Pred. No. 1.9e-24;
Matches 77; Conservative 35; Mismatches 92; Indels 9; Gaps 5;

Db 376 GTVCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKQKASLNGVNDSHMMASMLS 435
QY 178 --DGAAPRAGLAWSPCSRRLSLLSAGRACVWDPPR 213
Db 436 NLDHSQF-----WSPCSAYMTSFLDNGHGECLMDKPK 468

RESULT 12
US-09-963-791-16
; Sequence 16, Application US/09963791
; Patent No. 6649399
; GENERAL INFORMATION:
; APPLICANT: Donoho, Gregory
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Friedrich, Glenn
; APPLICANT: Scoville, John
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6649399el Human Proteases and Polynucleotides Encoding the Sa
; FILE REFERENCE: LEX-0105-USA
; CURRENT APPLICATION NUMBER: US/09/963,791
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/169,769
; PRIOR FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 317
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-16

Query Match 26.4%; Score 301; DB 4; Length 317;
Best Local Similarity 36.5%; Pred. No. 1.4e-24;
Matches 77; Conservative 32; Mismatches 78; Indels 24; Gaps 9;

QY 8 LELLVAGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 101 VETLVADKMMVGYHGRKDIEHYLSVMNIVAKLYRDSGLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLLSVCGWSQTI---NPEDDTDP-----GHADLVLYITRFDL----ELPD 113
Db 159 PNLEINHADKSLDSFCWKOKSILSHOSDGNTPENGIAHDNAVLITRYDICTYKNKPC 218
QY 114 GNRQVRGVTOLGACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPSGGPGSGH 173
Db 219 G----TGLASVAGCEPERSCSINEDIGLSAFTIAEIGHFNGVNHDDGI-GNSCGTKGH 274
QY 174 ---VNASDGAAPRAGLAWSPCSRRLSLL 200
Db 275 EAAKLMAAHITANTNPFSWSACS RDIYTSFL 305

Query Match 26.4%; Score 301; DB 4; Length 317;
Best Local Similarity 36.5%; Pred. No. 1.4e-24;
Matches 77; Conservative 32; Mismatches 78; Indels 24; Gaps 9;

QY 8 LELLVAGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 101 VETLVADKMMVGYHGRKDIEHYLSVMNIVAKLYRDSGLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLLSVCGWSQTI---NPEDDTDP-----GHADLVLYITRFDL----ELPD 113
Db 159 PNLEINHADKSLDSFCWKOKSILSHOSDGNTPENGIAHDNAVLITRYDICTYKNKPC 218
QY 114 GNRQVRGVTOLGACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPSGGPGSGH 173
Db 219 G----TGLASVAGCEPERSCSINEDIGLSAFTIAEIGHFNGVNHDDGI-GNSCGTKGH 274
QY 174 ---VNASDGAAPRAGLAWSPCSRRLSLL 200
Db 275 EAAKLMAAHITANTNPFSWSACS RDIYTSFL 305

Query Match 26.4%; Score 301; DB 4; Length 317;
Best Local Similarity 36.5%; Pred. No. 1.4e-24;
Matches 77; Conservative 32; Mismatches 78; Indels 24; Gaps 9;

QY 8 LELLVAGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 101 VETLVADKMMVGYHGRKDIEHYLSVMNIVAKLYRDSGLGNVNIIVARLIVLTEDQ-- 158
QY 67 PNITAN--LTSSLLSVCGWSQTI---NPEDDTDP-----GHADLVLYITRFDL----ELPD 113
Db 159 PNLEINHADKSLDSFCWKOKSILSHOSDGNTPENGIAHDNAVLITRYDICTYKNKPC 218
QY 114 GNRQVRGVTOLGACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPSGGPGSGH 173
Db 219 G----TGLASVAGCEPERSCSINEDIGLSAFTIAEIGHFNGVNHDDGI-GNSCGTKGH 274
QY 174 ---VNASDGAAPRAGLAWSPCSRRLSLL 200
Db 275 EAAKLMAAHITANTNPFSWSACS RDIYTSFL 305

; SEQ ID NO 6
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-963-791-6

Query Match 26.4%; Score 301; DB 4; Length 468;
Best Local Similarity 36.5%; Pred. No. 2.4e-24;
Matches 77; Conservative 32; Mismatches 78; Indels 24; Gaps 9;

QY 8 LELLVAGPDVFOAH-QEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 252 VETLVADKMMVGYHGRKDIEHYLSVMNIVAKLYRDSGLGNVNIIVARLIVLTEDQ-- 309
QY 67 PNITAN--LTSSLLSVCGWSQTI---NPEDDTDP-----GHADLVLYITRFDL----ELPD 113
Db 310 PNLEINHADKSLDSFCWKOKSILSHOSDGNTPENGIAHDNAVLITRYDICTYKNKPC 369
QY 114 GNRQVRGVTOLGACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPSGGPGSGH 173
Db 370 G----TGLASVAGCEPERSCSINEDIGLSAFTIAEIGHFNGVNHDDGI-GNSCGTKGH 425
QY 174 ---VNASDGAAPRAGLAWSPCSRRLSLL 200
Db 426 EAAKLMAAHITANTNPFSWSACS RDIYTSFL 456

RESULT 14
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; CURRENT FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 26.2%; Score 299.5; DB 4; Length 727;
Best Local Similarity 29.6%; Pred. No. 6.8e-24;
Matches 64; Conservative 36; Mismatches 101; Indels 15; Gaps 4;

QY 7 HLELLVAGPDVFOAHQEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 19 YVETMLVADQSMADFHGSLKHYLLTLFSAARFYKPSIRNSISLVVVKILVIYEEQKG 78
QY 67 PNITANLTSSLLSVCGWSQTIINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTLQGG 126
Db 79 PEVTSNAALTRNFCWSQKOHNSPDRDPEHYDTAILFTRODL-CGSHTCDTILGMADVGT 137
QY 127 ACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPG-----SGCGPSGHVMA----- 177
Db 138 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKHACASLNGVSGDLSHMLSSL 197
QY 178 DGAAPRAGLAWSPCSRRLSLLSAGRACVWDPPR 213
Db 198 DHSQP-----WSPCSAYMTSFLDNGHGECLMDKPK 228

Query Match 26.2%; Score 299.5; DB 4; Length 727;
Best Local Similarity 29.6%; Pred. No. 6.8e-24;
Matches 64; Conservative 36; Mismatches 101; Indels 15; Gaps 4;

QY 7 HLELLVAGPDVFOAHQEDTERVLTNLNIGAEILLRDPISLGAQFRVHLVKWVILTEPEGA 66
Db 19 YVETMLVADQSMADFHGSLKHYLLTLFSAARFYKPSIRNSISLVVVKILVIYEEQKG 78
QY 67 PNITANLTSSLLSVCGWSQTIINPEDDTPGHADLVLYITRFDLELPDGNRQVRGVTLQGG 126
Db 79 PEVTSNAALTRNFCWSQKOHNSPDRDPEHYDTAILFTRODL-CGSHTCDTILGMADVGT 137
QY 127 ACSPTWSCLTIEDTGFDLGVITIAEIGHSFGLHGDGAPG-----SGCGPSGHVMA----- 177
Db 138 VCDPSRSCSVIEDDGLQAAFTTAHELGHVFNPHDDAKHACASLNGVSGDLSHMLSSL 197
QY 178 DGAAPRAGLAWSPCSRRLSLLSAGRACVWDPPR 213
Db 198 DHSQP-----WSPCSAYMTSFLDNGHGECLMDKPK 228


```
RESULT 15
US-09-800-729-155
; Sequence 155, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-155

Query Match      25.9%; Score 295; DB 4; Length 2165;
Best Local Similarity 30.4%; Pred.No.1e-22;
Matches 70; Conservative 30; Mismatches 104; Indels 26; Gaps 3;

QY      7 HLELLVAVGPDVFQAHQEDTERVYLTNINIGAEILLRDPISLGAQFRVHLVXVILTEPEGA 66
DB      281 YEVLVVADTKMEYHGRSLEDYVLTTFSTVASIYRHSRLASINVVVKLIVLXTENAG 340

QY      67 PNITANLTSLLSVCGWSQTINPEDDDTPGHADLVLYITRFDLELPDGNRQVRGVTLGG 126
DB      341 PRITQAAQTLDQFCRWQYNDPDDSSVQHHVAILLTRKDICRSQKCDTLGLAELGT 400

QY      127 ACSPTWSCLITDPTGLGVITIAHIGHSFGLHGDGAPGSGCGP----- 170
DB      401 MCDMQXSALLIEDNGLSAAFTIAHELGHVFSIPHDDE--RKCSTYMPVNVKCKFQSTKFD 458

QY      171 -----SGHWASDGAAPRAGLAWSPCSRRLLSLL--SAGPARCVWDPP 212
DB      459 KTOFQNNFHIMAPTLEYNTHPWSWSPCSAGMLERFLENNRGQTCLFDQP 508
```

Search completed: March 13, 2004, 07:46:06
Job time : 10.0281 secs

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:33:19 ; Search time 2.4776 Seconds
(without alignments)
2756.533 Million cell updates/sec

Title: US-09-836-712-2_COPY_324_394

Perfect score: 399

Sequence: 1 PGLYISANEQCRVAFGPKAV.....VPLLDGTECGVEKWCCKGRC 71

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	33.8	1205	2 T18517	procollagen N-endo
2	124.5	31.2	508	2 T22836	hypothetical prote
3	121.5	30.5	951	2 T00017	gene ADAMTS-1 prot
4	119.5	29.9	550	2 T47158	hypothetical prote
5	111.1	27.8	2165	2 T21371	hypothetical prote
6	107.5	26.9	837	2 T00355	hypothetical prote
7	101.5	25.4	1444	2 T18856	angiogenesis inhib
8	78	19.5	610	2 JC7530	vascular apoptosis
9	74.5	18.7	860	2 T16892	hypothetical prote
10	73.5	18.4	957	2 T15976	hypothetical prote
11	66	16.5	826	2 A60385	monocyte surface a
12	65.5	16.4	471	2 B83127	probable secretion
13	65	16.3	384	2 S36529	E2 protein - human
14	64.5	16.2	145	2 S34574	gene 70 protein -
15	64	16.0	549	2 S48169	metalloproteinase
16	63.5	15.9	419	2 A59414	metalloproteinase
17	63	15.8	512	2 T37819	probable zinc meta
18	62.5	15.7	1042	2 T26644	hypothetical prote
19	62	15.5	357	2 S23403	sperm surface prot
20	62	15.5	425	2 C82775	phage-related prot
21	62	15.5	596	2 F87297	feruloyl-CoA synth
22	62	15.5	655	2 JC7850	disintegrin and me
23	62	15.5	1113	2 T80315	low-density lipopr
24	61.5	15.4	1620	2 T97283	hypothetical prote
25	61	15.3	64	2 A25775	metallothionein A
26	61	15.3	561	2 T06538	probable H-transp
27	61	15.3	584	1 C8HUA	complement C8 alph
28	60.5	15.2	360	2 S00834	int-1-like protein
29	60.5	15.2	651	2 T21175	hypothetical prote

30 60.5 15.2 2180 2 T29764 hypothetical prote
31 60 15.0 289 1 ICHUIB interleukin-1 beta
32 60 15.0 1286 2 A12195 two-component hybr
33 60 15.0 2470 2 I50726 cation-independent
34 59.5 14.9 571 2 S24789 jatarhagin C precu
35 59.5 14.9 1101 2 T16840 hypothetical prote
36 59 14.8 421 2 T96550 hypothetical prote
37 59 14.8 477 2 T25798 hypothetical prote
38 59 14.8 563 2 T32479 hypothetical prote
39 59 14.8 734 2 JC4861 ferritin beta cha
40 59 14.8 735 2 G02937 ferritin beta - cr
41 58.5 14.7 609 2 S55270 catrocollaetatin p
42 58.5 14.7 1144 2 D82521 hypothetical prote
43 58.5 14.7 1797 2 T21889 hypothetical prote
44 58.5 14.7 1805 2 T21888 hypothetical prote
45 58.5 14.7 3712 2 S18253 laminin alpha-1 ch

ALIGNMENTS

RESULT 1

T18517

procollagen N-endorpeptidase (EC 3.4.24.14) I - bovine

N;Alternate names: procollagen N-proteinase

C;Species: Bos primigenius taurus (cattle)

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T18517

R;Colligs, A.; Nuegens, B.V.; Lapiere, C.M.

submitted to the EMBL Data Library, February 1996

A;Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.

A;Reference number: Z18941

A;Accession: T18517

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-1205 <COL>

A;Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1

A;Experimental source: skin

C;Genetics:

A;Gene: PC I-NP

C;Function:

A;Description: catalyzes cleavage of the propeptides of type I and II collagens prior

C;Keywords: hydrolase; metalloproteinase

Query Match 33.8%; Score 135; DB 2; Length 1205;

Best Local Similarity 45.1%; Pred. No. 4.8e-07;

Matches 32; Conservative 5; Mismatches 32; Indels 2; Gaps 2;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDMQALSCHTDPLDQSCSRLLYPLLDGTEC 60

DB 476 PGLHYSMNEQCRFDGLGYMMCTAFRT-FDPCQLWC-SHPDNFYFCKTKKGPLDGTMC 533

QY 61 GVEKWCCKGRC 71

DB 534 APCKHCFKHC 544

RESULT 2

T22836

hypothetical protein F57B7.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C;Accession: T22836

R;Lennard, N.

submitted to the EMBL Data Library, June 1996

A;Reference number: Z19623

A;Accession: T22836

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-508 <WIL>

A;Cross-references: EMBL:Z74037; PIDN:CAA98493.1; GSPDB:GN00023; CESP:F57B7.4

A;Experimental source: clone F57B7

C;Genetics:

A;Gene: CESP:F57B7.4
A;Map position: 5
A;Introns: 45/3; 137/2; 221/2; 256/2; 306/2; 409/3; 451/3

Query Match 31.2%; Score 124.5; DB 2; Length 508;
Best Local Similarity 34.6%; Pred. No. 3.2e-06;
Matches 27; Conservative 8; Mismatches 30; Indels 13; Gaps 4;

QY 1 PGLYYSANECQVRVAFGPKAVACTFAREHLDQCALSCHTDPLDQSSCSRLVPLLDGTGC 53
Db 387 PGLVVKITRQCVAFGFTTIPCL---HIGYFHEQICERIMWCSGDSDE--CQTLNYPF 440

QY 54 LLDGTGCGVEKWCCKGRC 71
Db 441 AFDGTGCGYNNWCLGSC 458

RESULT 3
T00017
gene ADAMTS-1 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00017
R;Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A;Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A;Reference number: Z14055; MUID:98110583; PMID:9441751
A;Accession: T00017
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-951 <KUN>
A;Cross-references: EMBL:AB001735; NID:92809056; PIDN:BA24501.1; PID:92809057
A;Experimental source: strain 129SVJ
C;Genetics:
A;Gene: ADAMTS-1
A;Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
C;Superfamily: thrombospondin type 1 repeat homology
F;542-598/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 30.5%; Score 121.5; DB 2; Length 951;
Best Local Similarity 36.6%; Pred. No. 1.2e-05;
Matches 26; Conservative 4; Mismatches 38; Indels 3; Gaps 1;

QY 1 PGLYYSANECQVRVAFGPKAVACTFAREHLDQCALSCHTDPLDQSSCSRLVPLLDGTGC 60
Db 462 PGLYDANRQCQPTFGESKCPDA---ASTCTLTWCTGTGGLLVCQTKHFPWADGTSC 518

QY 61 GVEKWCCKGRC 71
Db 519 GEGKWCCKGRC 529

RESULT 4
T47158
hypothetical protein DKFP762C1110.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47158
R;Blum, H.; Bauersachs, S.; Meves, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24379
A;Accession: T47158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-550 <AAA>
A;Cross-references: EMBL:AL162080
A;Experimental source: adult melanoma (MeWo cell line); clone DKFP762C1110
C;Genetics:
A;Note: DKFP762C1110.1

Query Match 29.9%; Score 119.5; DB 2; Length 550;
Best Local Similarity 36.6%; Pred. No. 1.3e-05;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYYSANECQVRVAFGPKAVACTFAREHLDQCALSCHTDPLDQSSCSRLVPLLDGTGC 60
Db 61 PGLYDANRQCQPTFGESKCPDA---ASTCTLTWCTGTGGLLVCQTKHFPWADGTSC 117

QY 61 GVEKWCCKGRC 71
Db 118 GEGKWCCKGRC 128

RESULT 5
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C;Accession: T21371; T24896
R;Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19413
A;Accession: T21371
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2165 <WIL>
A;Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A;Experimental source: clone F25H8
R;Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A;Reference number: Z19949
A;Accession: T24896
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-2165 <WIL>
A;Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A;Experimental source: clone T13H10
C;Genetics:
A;Gene: CESP:F25H8.3
A;Map position: 4
A;Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 1

Query Match 27.8%; Score 111; DB 2; Length 2165;
Best Local Similarity 32.9%; Pred. No. 0.00039;
Matches 24; Conservative 10; Mismatches 33; Indels 6; Gaps 2;

QY 1 PGLYYSANECQVRVAFGPKAVACTFAREHLDQCALSCHTDPLDQSSCSRLVPLLDGTGC 60
Db 523 PGKXDAHQCKVFVGFPASELCPY---MPTCRRLWCATFYGSQMGCRTHMFWADGTSC 578

QY 61 GVEK--WCSKGR 71
Db 579 DBSRSMFCHHGAC 591

RESULT 6
T00355
hypothetical protein KIAA0688 - human
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00355
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, I.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complex
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z14142; MUID:98403880; PMID:9734811
A;Accession: T00355
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-837 <ISH>
A;Cross-references: EMBL:AB014568; NID:93327189; PIDN:BA31663.1; PID:93327190
A;Experimental source: brain
C;Genetics:
A;Gene: KIAA0688
C;Superfamily: thrombospondin type 1 repeat homology
F;519-575/Domain: thrombospondin type 1 repeat homology <THR3>

Query Match 26.9%; Score 107.5; DB 2; Length 837;
Best Local Similarity 35.2%; Pred. No. 0.0004;
Matches 25; Conservative 7; Mismatches 36; Indels 3; Gaps 1;

QY 1 PGLYSANQCRVAFQPKVACTFAREHLDMCOALSCHTDPLDQSSCSRLVPLLDGTGEC 60
DB 439 PKKDADADQCTQTFDPSRHP---QLPFCALMCSGHLNGHAMCOTKESPWADGTPC 495

QY 61 GVEKWCCKGRC 71
DB 496 GPAQACMGGRG 506

RESULT 7
T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18856; T24653
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: T18856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL>
A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GNO0028; CESP:CO2B4.1
A:Experimental source: clone C0284
R:McMurray, A.
submitted to the EMBL Data Library, July 1995
A:Reference number: Z19917
A:Accession: T24653
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <W12>
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GNO0028; CESP:CO2B4.1
A:Experimental source: clone T07C5
C:Genetics:
A:Gene: CESP:CO2B4.1
A:Map position: X
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 568

Query Match 25.4%; Score 101.5; DB 2; Length 1444;
Best Local Similarity 30.8%; Pred. No. 0.003;
Matches 24; Conservative 11; Mismatches 26; Indels 17; Gaps 3;

QY 1 PGLYSANQCRVAFG-----PKVACTFAREHLDMCOALSCHTDPLDQSSCSRLVLP 53
DB 449 PQCRFTADQCCSYFWRGDKYKVEIPNGKAMD-----DICRILWCG-----NSGSTISTAHP 498

QY 54 LLDGTGCGVEKWCCKGRC 71
DB 499 ALEGSWCGANKWCHKGQC 516

RESULT 8
JC7530
vascular apoptosis-inducing protein 1 - western diamondback rattlesnake
C:Species: Crotalus atrox (western diamondback rattlesnake)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 01-Mar-2002
C:Accession: JC7530; PC7105
R:Masuda, S.; Ohta, T.; Kaji, K.; Fox, J.W.; Hayashi, H.; Araki, S.
Biochem. Biophys. Res. Commun. 278, 197-204, 2000
A:Title: cDNA cloning and characterization of vascular apoptosis-inducing protein 1.
A:Reference number: JC7530; MUID:20525424; PMID:11071872
A:Contents: Crude venom gland
A:Accession: JC7530
A:Molecule type: mRNA
A:Residues: 1-610 <MAS>
A:Cross-references: DBJ:AB042840
A:Accession: PC7105
A:Molecule type: protein
A:Residues: 240-251;274-283;378-385;507-519 <MA2>

C:Comment: This protein, as a homodimeric protein, belongs to snake venom metalloprotease a specific substrate, induces apoptosis in vascular endothelial cells.
C:Genetics:
A:Gene: vapi
C:Superfamily: mouse meltrin alpha; disintegrin homology
C:Keywords: apoptosis

Query Match 19.5%; Score 78; DB 2; Length 610;
Best Local Similarity 27.4%; Pred. No. 0.57;
Matches 26; Conservative 11; Mismatches 26; Indels 32; Gaps 6;

QY 9 EOCRVAFGPKAV---AC-TFARF--HLMCQ-----ALSCH-----TDPLD 43
DB 509 DQCIALFGGATVYQDACQFQNEGNYGYCKRKQNTKIACEPQDKGRLYCFNNSPEN 568

QY 44 QSSCSRLVLP-----LLDGTGCGVEKWCCKGRC 71
DB 569 KMPCNTYVSPNDEDKGMVLFSTKCADRKACSNQGC 603

RESULT 9
T16892
hypothetical protein T19D2.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16892
R:Bentley, D.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid T19D2.
A:Reference number: Z18599
A:Accession: T16892
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-860 <SEN>
A:Cross-references: EMBL:U42846; NID:g1125809; PID:g1125810; PIDN:AAA83600.1; CESP:T191
C:Genetics:
A:Gene: CESP:T19D2.1
A:Introns: 56/3; 96/1; 171/3; 199/3; 245/2; 293/2; 335/3; 395/3; 426/3; 484/2; 505/3; 5

Query Match 18.7%; Score 74.5; DB 2; Length 860;
Best Local Similarity 27.3%; Pred. No. 1.9;
Matches 21; Conservative 11; Mismatches 30; Indels 15; Gaps 3;

QY 1 PGLYSANQCRVAFQPKVACTFAREHLD-----MQALSCHTDPLDQSSCSRLVLP 54
DB 366 PGQLFLDQECEIFHGE-----CWKHELKQGTVMQICQMVWCG---NGEGVIRTAHPA 416

QY 55 LDGTGCGVEKWCCKGRC 71
DB 417 LEGTCGFGMICRQGC 433

RESULT 10
T15976
hypothetical protein F08C6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T15976
R:Bentley, D.
submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmid F08C6.
A:Reference number: Z18440
A:Accession: T15976
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-957 <SEN>
A:Cross-references: EMBL:U29378; NID:9868184; PID:9868185; PIDN:AAA68721.1; CESP:F08C6
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F08C6.1
A:Introns: 23/1; 135/3; 220/3; 285/3; 325/2; 376/3; 461/2; 498/3; 532/3; 627/2; 714/1;

Query Match 18.4%; Score 73.5; DB 2; Length 957;

Best Local Similarity 38.6%; Pred. No. 2.7;
Matches 22; Conservative 3; Mismatches 19; Indels 13; Gaps 4;
QY 22 CTFAHEHLMCOA-LSCH-----TLPDQSSCSRLVPLLDGTECGVEKWCCKGRC 71
Db 353 CTLA-EGLDFTSAFIGTHLGHRTXDKHSDT-----LFTGFCGSKWCQIGRC 403
RESULT 11
A60385
monocyte surface antigen MS2 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 26-Aug-1999
C:Accession: A60385
R.Yoshida, S.; Setoguchi, M.; Higuchi, Y.; Akizuki, S.; Yamamoto, S.
Int. Immunol. 2, 585-591, 1990
A:Title: Molecular cloning of cDNA encoding MS2 antigen, a novel cell surface antigen st
A:Reference number: A60385; MUID:91197896; PMID:1982220
A:Accession: A60385
A:Molecule type: mRNA
A:Residues: 1-826 <YOS>
A:Cross-references: EMBL:X13335
C:Superfamily: mouse mectrin alpha; disintegrin homology
C:Keywords: glycoprotein; surface antigen; transmembrane protein
F:1-14/Domain: signal sequence #status predicted <SIG>
F:402-484/Domain: disintegrin homology <DIS>
F:659-683/Domain: transmembrane #status predicted <TM>
F:7330/Active site: Glu #status predicted
Query Match 16.5%; Score 66; DB 2; Length 826;
Best Local Similarity 25.8%; Pred. No. 16;
Matches 25; Conservative 11; Mismatches 27; Indels 34; Gaps 5;
QY 9 EGCRAVAFGP-----KAVAC-TFA-----REHLMQALSCH--TDPDQSSCS--- 48
Db 508 QQCRLWGPARGVAADSCYTFIPPGNGRMYSGRINRCGALYCBGGQKPLERSFCTFS 567
QY 49 -----RLVPLLDGTECGVEKWCCKGRC 71
Db 568 NHGCVHALGTSGNIDTPELVLOQTKEEGKVCMDGSC 604
RESULT 12
B83127
probable secretion protein PA4144 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83127
R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.; Bu
adnan, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 405, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: B83127
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-471 <STO>
A:Cross-references: GB:AE004831; GB:AE004091; NID:9950347; PIDN:AG07531.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4144
Query Match 16.4%; Score 65.5; DB 2; Length 471;
Best Local Similarity 34.6%; Pred. No. 11;
Matches 18; Conservative 8; Mismatches 25; Indels 1; Gaps 1;
QY 9 EGCRAVAFGP-
Db 304 EESRAGRPSIALSANLARSHSDQAVAFNGDTRDRSRIGLQNLPLFEGFE 355
RESULT 13

S36529
E2 protein - human papillomavirus type 53
C:Species: human papillomavirus type 53
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36529
R.Delius, H.; Hofmann, B.
Submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36529
A:Molecule type: DNA
A:Residues: 1-384
A:Cross-references: EMBL:X74482; NID:9397046; PIDN:CAA52593.1; PID:9397049
C:Superfamily: papillomavirus E2 protein
C:Keywords: DNA binding; early protein; transcription regulation
Query Match 16.3%; Score 65; DB 2; Length 384;
Best Local Similarity 26.4%; Pred. No. 10;
Matches 19; Conservative 9; Mismatches 26; Indels 18; Gaps 3;
QY 3 LYTSANEQCRVAFGPKAVACTFAREHLMQALSCHTDPDLD---QSSCSRLVPLLDGTE 59
Db 42 IYKARENNTKLGHVQVPC-----LQVKAKACVAIELOIALESCK-----TE 86
QY 60 CGVEKWCCKGRC 71
Db 87 YNMEWTLRDVC 98
RESULT 14
S34574
gene 70 protein - Mycobacterium phage L5
C:Species: Mycobacterium phage L5
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Jun-2000
C:Accession: S34574; S31015
R.Hatfull, G.F.
Submitted to the EMBL Data Library, December 1992
A:Reference number: S34574
A:Accession: S34574
A:Molecule type: DNA
A:Residues: 1-145 <HAT>
A:Cross-references: EMBL:Z18946; NID:915859; PIDN:CAA79446.1; PID:G579144
R.Donnelly-Wu, M.K.; Jacobs Jr., W.R.; Hatfull, G.F.
MOJ. Microbiol. 7, 407-417, 1993
A:Title: Superinfection immunity of mycobacteriophage L5: applications for genetic tra
A:Reference number: S30949; MUID:93211283; PMID:8459767
A:Accession: S31015
A:Molecule type: DNA
A:Residues: 1-81 <DON>
A:Cross-references: EMBL:Z18946
A:Note: the authors translated the initiation codon GTG for residue 1 as Val
C:Genetics:
A:Gene: 70
A:Start codon: GTG
C:Superfamily: Mycobacterium phage L5 gene 70 protein
Query Match 16.2%; Score 64.5; DB 2; Length 145;
Best Local Similarity 33.8%; Pred. No. 4.9;
Matches 25; Conservative 8; Mismatches 32; Indels 9; Gaps 3;
QY 3 LYTSANEQCRVA-----FGPKAVACTFAREH--LMQALSCHTDPDLDQSSCSRLVPLLD 56
Db 71 LILSANKVEYVARGTFRDPTMIHCDKGTWVLRMTQADQFPKDPIDMSEVAKLLSVAVD 130
QY 57 GTECGVEKWCCKGRC 70
Db 131 G---ALNETTGKGR 141
RESULT 15
S48169
metalloprotease (EC 3.4.24.-) H-II precursor - carpet viper (fragment)
N;Contains: disintegrin

C:Species: Echis pyramidum leakeyi
C:Date: 06-Dec-1996 #sequence_revision 06-Dec-1996 #text_change 09-Jun-2000
C:Accession: S48169
R:Paire, M.J.I.; Moura-Da-Silva, A.M.; Theakston, R.D.G.; Crampton, J.M.
Eur. J. Biochem. 224, 483-488, 1994
A:Title: Cloning of metalloprotease genes in the carpet viper (Echis pyramidum leakeyi).
A:Reference number: S48169; MUID:95010025; PMID:7925363
A:Accession: S48169
A:Molecule type: mRNA
A:Residues: 1-549 <PAI>
A:Cross-references: EMBL:X78971; NID:g763094; PIDN:CAA55566.1; PID:g763095
C:Superfamily: mouse melitin alpha; disintegrin homology
C:Keywords: hydrolase; metalloprotease; venom; zinc
F:1-138/Domain: propeptide (fragment) #status predicted <PRO>
F:139-549/Product: metalloprotease H-II #status predicted <MAT>
F:350-432/Domain: disintegrin homology <DIS>
F:284/Active site: Glu #status predicted

Query Match 16.0%; Score 64; DB 2; Length 549;
Best Local Similarity 27.2%; Pred. No. 19;
Matches 25; Conservative 8; Mismatches 23; Indels 36; Gaps 6;

Qy 10 QCRVAFGPKAV---AC-----TFAREHLD-----COALSCHTDPLDQS 45
Db 457 QCYAHFGQNAVVGQDACPEINKEGKGFYCRKENDVIPCQEDVKCGRLFCETEP---N 513

Qy 46 SCSRLLVPLLD-----GTECGVEKWCCKGRC 71
Db 514 MCR---YPYDGEWVDPGTYCEDKVCVINGKC 542

Search completed: March 13, 2004, 07:44:36
Job time : 3.4776 secs

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 1.47917 Seconds
(without alignments)
2499.367 Million cell updates/sec

Title: US-09-836-712-2_COPY_324_394

Perfect score: 399

Sequence: 1 POLYISANEQCRVAFGPKAV.....VPLLDGTGCGVKWCSKGR 71

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	181	45.4	1095	1	AT17_HUMAN
2	148	37.1	1207	1	Q8te56 homo sapien
3	146	36.6	1210	1	AT19_MOUSE
4	144	36.1	1205	1	O15072 mus musculus
5	143.5	36.0	1593	1	O15072 homo sapien
6	136	34.1	1081	1	Q8te60 homo sapien
7	135	33.8	1205	1	P79331 b adants-2
8	135	33.8	1211	1	O95450 h adants-2
9	135	33.8	1213	1	Q8c9w3 m adants-2
10	135	33.8	1224	1	Q8te57 homo sapien
11	134	33.6	1223	1	Q8wx88 homo sapien
12	128	32.1	1911	1	P59510 homo sapien
13	124.5	31.2	967	1	Q9wuq1 rattus norv
14	124.5	31.2	997	1	Q8ukp4 homo sapien
15	124	31.1	562	1	P59384 mus musculus
16	123	30.8	207	1	Q8tt92 bos taurus
17	122	30.6	930	1	Q9una0 homo sapien
18	122	30.6	930	1	Q9r001 mus musculus
19	122	30.6	1906	1	P59511 mus musculus
20	121.5	30.5	968	1	P7857 mus musculus
21	119.5	29.9	967	1	Q9uh18 homo sapien
22	119	29.8	1935	1	Q8p2n4 homo sapien
23	118	29.6	950	1	Q8te58 homo sapien
24	116	29.1	890	1	Q8up79 homo sapien
25	108.5	27.2	630	1	Q8esp7 rattus norv
26	107.5	26.9	245	1	Q8tt93 bos taurus
27	107.5	26.9	837	1	O75173 homo sapien
28	106.5	26.7	824	1	P78325 homo sapien
29	96.5	24.2	1077	1	Q9h324 homo sapien
30	77.5	19.4	505	1	P57110 mus musculus
31	72.5	18.2	445	1	P97776 rattus norv
32	72.5	18.2	719	1	Q9r157 mus musculus
33	69	17.3	405	1	Q8arl1 vibrio vuln

34	66.5	16.7	860	1	AT56_HUMAN	Q9ukp5 homo sapien
35	66	16.5	826	1	AD08_MOUSE	Q05910 mus musculus
36	65.5	16.4	405	1	GLC1_VIBPA	Q87qx6 vibrio para
37	65	16.3	384	1	VE2_FPV53	P36797 human papil
38	64.5	16.2	145	1	VG70_BPML5	Q05284 mycobacteri
39	63	15.8	401	1	KACU_NAJKA	P82942 naja kaouth
40	63	15.8	512	1	YE94_SCHPO	O13766 schizosacch
41	62	15.5	735	1	AD02_CAVPO	Q68411 cavia porce
42	62	15.5	819	1	AD09_HUMAN	Q13443 homo sapien
43	62	15.5	1113	1	CORI_MOUSE	Q92319 mus musculus
44	61	15.3	584	1	CO8A_HUMAN	P07357 homo sapien
45	61	15.3	833	1	SRC2_MOUSE	P59222 mus musculus

ALIGNMENTS

RESULT 1

AT17_HUMAN

ID AT17_HUMAN STANDARD; PRT; 1095 AA.

AC Q8TE56;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-17 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase

DE with thrombospondin motifs 17) (ADAM-TS 17) (ADAM-TS17).

GN ADAMTS17.

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21856482; PubMed=11867212;

RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,

RA Lopez-Otin C.;

RT "Cloning, expression analysis, and structural characterization of

RT seven novel human ADAMTS, a family of metalloproteinases with

RT disintegrin and thrombospondin-1 domains.";

RL Gene 283:49-62(2002).

CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular

CC matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, in adult brain,

CC prostate, and liver.

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By

CC similarity).

CC -!- SIMILARITY: Belongs to peptidase family M12B.

CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC -!- SIMILARITY: Contains 1 PLAC domain.

CC -!- SIMILARITY: Contains 5 TSP type-1 domains.

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AJ315735; CAC86016.1; -

DR MEROPS; M12.027; -

DR Genew; HGNC:17109; ADAMTS17.

DR MIM; 607511; -

DR InterPro; IPR001762; Disintegrin.

DR InterPro; IPR001818; Pept_M10A_M12B.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR001590; Peptidase_M12B.

DR InterPro; IPR002870; Peptidase_M12B_N.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.

DR Pfam; PF01562; Pep_M12B_Proprep; 1.

DR Pfam; PF01421; Reprolysin; 1.

DR Pfam; PF00090; tsp_1; 5.

DR PRINTS; PR01705; TSPIREPEAT.
 DR SMART; SMO0209; TSPI; 5.
 DR PROSITE; PS00215; ADAM MEPRO; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; FALSE NEG.
 DR PROSITE; PS00900; PLAC; 1.
 DR PROSITE; PS00092; TSPI; 5.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 FT Repeat; Extracellular matrix. POTENTIAL.
 FT SIGNAL 1 27
 FT PROPEP 28 223
 FT CHAIN 224 1095
 FT DOMAIN 224 452
 FT DOMAIN 453 542
 FT DOMAIN 543 598
 FT DOMAIN 599 701
 FT DOMAIN 702 779
 FT DOMAIN 800 860
 FT DOMAIN 861 922
 FT DOMAIN 925 968
 FT DOMAIN 972 1029
 FT DOMAIN 1045 1084
 FT DOMAIN 60 120
 FT SITE 201 201
 FT METAL 389 389
 FT ACT SITE 390 390
 FT METAL 393 393
 FT METAL 399 399
 FT CARBOHYD 167 167
 FT CARBOHYD 483 483
 FT CARBOHYD 785 785
 FT CARBOHYD 790 790
 FT CARBOHYD 832 832
 FT CARBOHYD 839 839
 FT CARBOHYD 894 894
 FT SEQUENCE 1095 AA; 121099 MW; A5C6540484754D5D CRC64;
 Query Match 45.4%; Score 181; DB 1; Length 1095;
 Best Local Similarity 52.1%; Pred. No. 6.7e-13;
 Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;
 QY 1 PGLYSANQCRVARGPKVACTFAR--EHLDMQALSCHTDPDQSSCSRLVPLLDGT 58
 Db 466 PGMHYSANEQCQLIFGMNA---TFCRNWEHL-MCAGLMCLVE--GDTSCKTKLDPLDGT 519
 QY 59 ECGVEKWCCKGRC 71
 Db 520 ECGADKWCRCAGEC 532
 RESULT 2
 AT19_HUMAN STANDARD; PRT; 1207 AA.
 ID AT19_HUMAN
 AC Q8TE59;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 with thrombospondin motifs 19) (ADAM-TS 19) (ADAM-TS19).
 GN ADAMTS19.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 seven novel human ADAMTSs, a family of metalloproteinases with

disintegrin and thrombospondin-1 domains.";
 Gene 283:49-62(2002).
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, but not in any adult
 tissues examined. Expression was detected in an osteosarcoma cDNA
 library.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; AJ311904; CAC84565.1; .
 Genew; HGNC:17111; ADAMTS19.
 MIM; 607513; .
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR001818; Pept_M12A_M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR InterPro; IPR000884; TSPI.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; RepPolysin; 1.
 DR Pfam; PF00090; tsp; 1; 5.
 DR SMART; SMO0209; TSPI; 5.
 DR PROSITE; PS00215; ADAM MEPRO; 1.
 DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; FALSE NEG.
 DR PROSITE; PS00900; PLAC; 1.
 DR PROSITE; PS00092; TSPI; 5.
 DR PROSITE; PS00142; ZINC PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix. POTENTIAL.
 KW SIGNAL 1 21
 KW PROPEP 22 316
 KW CHAIN 317 1207
 KW DOMAIN 317 545
 KW DOMAIN 546 633
 KW DOMAIN 634 686
 KW DOMAIN 687 790
 KW DOMAIN 791 914
 KW DOMAIN 915 975
 KW DOMAIN 976 1037
 KW DOMAIN 1039 1083
 KW DOMAIN 1087 1144
 KW DOMAIN 1160 1199
 KW DOMAIN 88 216
 KW SITE 294 294
 KW METAL 482 482
 KW ACT SITE 483 483
 KW METAL 486 486
 KW METAL 492 492
 KW CARBOHYD 260 260
 KW CARBOHYD 797 797
 KW CARBOHYD 907 907
 KW CARBOHYD 949 949
 KW CARBOHYD 1009 1009
 KW SEQUENCE 1207 AA; 134061 MW; AF36F6BF586FDE2 CRC64;
 Query Match 37.1%; Score 148; DB 1; Length 1207;
 Best Local Similarity 39.4%; Pred. No. 4.3e-09;

Matches 28; Conservative 13; Mismatches 26; Indels 4; Gaps 3;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLQSSCSRLVPLLDGTTC 60
 DB 559 PGMYTADQCQILFGPLASFCQ-EMQHV-ICTGLWCXVE--GEACRTKLDPPMDGTDC 614
 QY 61 GVEKWCCKGRC 71
 DB 615 DLGKCKAGEC 625

AT19_MOUSE
 ID AT19_MOUSE STANDARD; PRT; 1210 AA.
 AC P59509;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 19) (ADAM-TS 19) (ADAM-TS19).
 GN ADAMTS19.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Ovary;
 RX MEDLINE=22505168; PubMed=12617826;
 RA Menke D.B., Page D.C.;
 RT "Sexually dimorphic gene expression in the developing mouse gonad.";
 RL Gene Expr. Patterns 2:359-367(2002).
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in fetal ovary, low
 CC levels of expression is also detected in kidney, heart, skeletal
 CC muscle, lung and testis.
 CC -!- DEVELOPMENTAL STAGE: Expression is strongest in anterior and
 CC ventral regions of the ovary at 12.5 and 13.5 dpc before becoming
 CC more uniform.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 5 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC -!- CAUTION: By homology with the human sequence, it is uncertain
 CC whether Met-1 or Met-5 is the initiator.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AY135183; AAN10155.1; .
 DR MGD; MGI:2442875; Adamts19.
 DR InterPro; IPR006586; ADAM cysteine.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR001818; Pept M10A_M12B.
 DR InterPro; IPR006025; Pept M_Zn_BS.
 DR InterPro; IPR001590; Peptidase M12B.
 DR InterPro; IPR002870; Peptidase M12B_N.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; tsp 1; 5.
 DR SMART; SMO0608; ACF 1.
 DR SMART; SMO0209; TSP1; 5.
 DR PROSITE; PS50215; ADAM_MEPRO; 1.
 DR

DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
 DR PROSITE; PS00427; DISINTEGRIN 1; FALSE NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS00900; PLAC; 1.
 DR PROSITE; PS00092; TSPL; 5.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix.
 FT SIGNAL 1 30 POTENTIAL.
 FT PROPEP 31 319 BY SIMILARITY.
 FT CHAIN 320 1210 ADAMTS-19.
 FT DOMAIN 320 548 METALLOPROTEASE.
 FT DOMAIN 549 636 DISINTEGRIN-LIKE.
 FT DOMAIN 637 689 TSP TYPE-1 1.
 FT DOMAIN 690 793 CYS-RICH.
 FT DOMAIN 794 917 SPACER.
 FT DOMAIN 918 978 TSP TYPE-1 2.
 FT DOMAIN 979 1040 TSP TYPE-1 3.
 FT DOMAIN 1042 1086 TSP TYPE-1 4.
 FT DOMAIN 1090 1147 TSP TYPE-1 5.
 FT DOMAIN 1163 1202 PLAC.
 FT SITE 297 297 CYSTEINE SWITCH (POTENTIAL).
 FT METAL 485 485 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT_SITE 486 486 BY SIMILARITY.
 FT METAL 489 489 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 495 495 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 800 800 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 910 910 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1012 1012 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1210 AA; 134560 MW; 0AB812ABAB4BB7A2 CRC64;
 Query Match 36.8%; Score 146; DB 1; Length 1210;
 Best Local Similarity 39.4%; Pred. No. 7.3e-09;
 Matches 28; Conservative 13; Mismatches 26; Indels 4; Gaps 3;
 QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLQSSCSRLVPLLDGTTC 60
 DB 562 PGMYTADQCQILFGPLASFCQ-EMQHV-ICTGLWCXVE--GEACRTKLDPPMDGTDC 617
 QY 61 GVEKWCCKGRC 71
 DB 618 DPGKCKAGEC 628

RESULT 4
 AT33 HUMAN
 ID AT33 HUMAN STANDARD; PRT; 1205 AA.
 AC G15072; Q9BXZ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 3) (ADAM-TS 3) (ADAM-TS3) (Procollagen II
 DE amino-propeptide processing enzyme) (Procollagen II N-proteinase) (PC
 DE II-NP).
 DE GN ADAMTS3 OR KIAA0366.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-227 FROM N.A.
 RX MEDLINE=21402912; PubMed=11408482;
 RA Fernandes R.J., Hirohata S., Engle J.M., Colige A., Cohn D.H.,
 RA Eyre D.R., Apte S.S.;
 RT "Procollagen II amino propeptide processing by ADAMTS-3. Insights on
 RT dermatosparaxis.";
 RL J. Biol. Chem. 276:31502-31509(2001).
 RN [2]

SEQUENCE OF 5-1205 FROM N.A.
TISSUE=Brain;
MEDLINE=97349984; PubMed=9205841;
Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
"Prediction of the coding sequences of unidentified human genes. VII.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro."
DNA Res. 4:141-150(1997).
-!- FUNCTION: Cleaves the propeptides of type II collagen prior to
fibril assembly. Does not act on types I and III collagens.
-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
-!- TISSUE SPECIFICITY: Found in cartilage and skin.
-!- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix.
-!- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
-!- SIMILARITY: Belongs to peptidase family M12B.
-!- SIMILARITY: Contains 1 disintegrin-like domain.
-!- SIMILARITY: Contains 1 PLAC domain.
-!- CAUTION: Has sometimes been referred to as ADAMTS4.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF247668; AAX28400.1; -;
EMBL; AB002364; BAA20821.1; -;
MEROPS; M12.220; -;
Genew; HGNC:219; ADAMTS3.
MIM; 605011; -;
GO; GO:0005578; C:extracellular matrix; NAS.
GO; GO:0004222; F:metalloendopeptidase activity; NAS.
GO; GO:0008270; F:zinc ion binding; NAS.
GO; GO:0030574; P:collagen catabolism; NAS.
GO; GO:0030199; P:collagen fibril organization; NAS.
InterPro; IPR001762; Disintegrin.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR001590; Peptidase_M12B.
InterPro; IPR002870; Peptidase_M12B_N.
InterPro; IPR000894; TSP1.
InterPro; IPR008085; TSP 1.
Pfam; PF01562; Pfam_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; tsp 1; 4.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1; 4.
PROSITE; PS50215; ADAM_MEPRO; 1; FALSE_NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS50900; PLAC; 1.
PROSITE; PS50092; TSP1; 4.
PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydroxylase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 20
FT PROPEP 21 249
FT CHAIN 250 1205
FT DOMAIN 250 469
FT DOMAIN 470 550
FT DOMAIN 551 606
FT DOMAIN 608 712
FT DOMAIN 713 844
FT DOMAIN 845 905
FT DOMAIN 906 965
FT DOMAIN 966 1014

FT DOMAIN 1015 1054
FT DOMAIN 246 249
FT METAL 398 398
FT ACT SITE 399 399
FT METAL 402 402
FT METAL 408 408
FT CARBOHYD 83 83
FT CARBOHYD 119 119
FT CARBOHYD 242 242
FT CARBOHYD 345 345
FT CARBOHYD 475 475
FT CARBOHYD 814 814
FT CARBOHYD 942 942
SQ SEQUENCE 1205 AA; 135570 MW; EB07B286FC85FB87 CRC64;
Query Match 36.1%; Score 144; DB 1; Length 1205;
Best Local Similarity 45.1%; Pred. No. 1.2e-08;
Matches 32; Conservative 5; Mismatches 32; Indels 2; Gaps 2;
QY 1 POLYISANEQCRVARGPKAVACTFAHEHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 472 PGINYSMDQCRDFGVGYMCTAFRT-FDPCQLMC-SHPDNPYFKTKGKGPLDGTGTEC 529
QY 61 GVEKWCCKGRC 71
Db 530 AAGKWCYKRGHC 540
RESULT 5
AT12 HUMAN STANDARD; PRT; 1593 AA.
ID AT12 HUMAN
AC P58397;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
GN ADAMTS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=21264577; PubMed=11279086;
RA Cal S., Argueelles J.M., Fernandez P.L., Lopez-Otin C.;
"Identification, characterization, and intracellular processing of
RT ADAM-TS12, a novel human disintegrin with a complex structural
RT organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940(2001).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC expressed in gastric carcinomas and in cancer cells of diverse
CC origin.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase.
CC -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING
CC TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE
CC METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND
CC THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1
CC DOMAINS
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 8 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ250725; CAC20419.1; -
Genew; HGNC:14605; ADAMTS12.
MIM; 606184; -
InterPro; IPR001762; Disintegrin.
InterPro; IPR001818; Pept_M10A_M12B.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR001590; Peptidase_M12B.
InterPro; IPR002870; Peptidase_M12B_N.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP1.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; TSP1; 6.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1; 8.
PROSITE; PS50215; ADAM_MEPRO; 1.
PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS00900; PLAC; 1.
PROSITE; PS0092; TSP1; 6.
PROSITE; PS00142; ZINC_PROTEASE; 1.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Extracellular matrix.
SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 240 BY SIMILARITY.
FT CHAIN 241 1593 ADAMTS-12.
FT DOMAIN 241 464 METALLOPROTEASE.
FT DOMAIN 465 544 DISINTEGRIN-LIKE.
FT DOMAIN 542 597 TSP TYPE-1 1.
FT DOMAIN 597 700 CYS-RICH.
FT DOMAIN 701 826 SPACER 1.
FT DOMAIN 823 882 TSP TYPE-1 2.
FT DOMAIN 886 942 TSP TYPE-1 3.
FT DOMAIN 943 996 TSP TYPE-1 4.
FT DOMAIN 996 1315 SPACER 2.
FT DOMAIN 1312 1365 TSP TYPE-1 5.
FT DOMAIN 1367 1421 TSP TYPE-1 6.
FT DOMAIN 1422 1470 TSP TYPE-1 7.
FT DOMAIN 1471 1531 TSP TYPE-1 8.
FT DOMAIN 1534 1574 PLAC.
FT DOMAIN 302 305 POLY-GLU.
FT SITE 208 208 CYSTEINE_SWITCH (POTENTIAL).
FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 393 393 BY SIMILARITY.
FT METAL 396 396 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1300 1300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1371 1371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1503 1503 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1593 AA; 177545 MW; 07F9F48E63BD83A3 CRC64;

Query Match 36.0%; Score 143.5; DB 1; Length 1593;
Best Local Similarity 38.0%; Pred. No. 1.9e-08;
Matches 27; Conservative 12; Mismatches 25; Indels 7; Gaps 2;

QY 1 PGLYIANEQRVAFGPKAVACTFAREHLMQCALSCHTDPDQSSCRLLVPLDGTGC 60
Db 468 PGVIYVHHQCQIQYGNNA---TFCQEVNVCQTLWCSV----KGFCRSKLDAAADGTQC 520
QY 61 GVEKWCSGKRC 71
Db 521 GEKKWCWAGKC 531

RESULT 6
AT18_HUMAN
ID AT18_HUMAN STANDARD; PRT; 1081 AA.
AC Q8TE60;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-18 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 18) (ADAM-TS 18) (ADAM-TS18).
GN ADAMTS18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=1187212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V., Lopez-Otin C.;
RT Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney and in adult brain, prostate, submaxillary gland, and endothelium.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ311903; CAC83612.1; -
Genew; HGNC:17110; ADAMTS18.
MIM; 607512; -
InterPro; IPR001762; Disintegrin.
InterPro; IPR001818; Pept_M10A_M12B.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR001590; Peptidase_M12B.
InterPro; IPR002870; Peptidase_M12B_N.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP1.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; TSP1; 4.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1; 4.
PROSITE; PS50215; ADAM_MEPRO; 1.
PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS00900; TSP1; 3.
PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

KW Repeat; Extracellular matrix.

FT SIGNAL 1 47

FT PROPEP 48 284

FT CHAIN 285 1081

FT DOMAIN 285 497

FT DOMAIN 498 577

FT DOMAIN 588 643

FT DOMAIN 644 749

FT DOMAIN 750 876

FT DOMAIN 877 931

FT DOMAIN 933 991

FT DOMAIN 993 1048

FT METAL 254 254

FT METAL 435 435

FT ACT_SITE 436 436

FT METAL 439 439

FT METAL 445 445

FT CARBOHYD 151 151

FT CARBOHYD 190 190

FT CARBOHYD 313 313

FT CARBOHYD 744 744

FT CARBOHYD 837 837

FT CARBOHYD 908 908

SQ SEQUENCE 1081 AA; 119655 MW; 0438BF645676461E CRC64;

Query Match 34.1%; Score 136; DB 1; Length 1081;

Best Local Similarity 37.0%; Pred. No. 8.9e-08;

Matches 27; Conservative 12; Mismatches 26; Indels 8; Gaps 3;

QY 1 PGLYSANEQCVAPGKAVACT--FAREHLDMCALSCHTDPLDQSSCSRLVPLDGT 58

DB 510 PGOIYDADTQWQFGAKAKLCSLGFK---DICKSLWCHR---VGHRCETKFPAAEGT 563

QY 59 ECGVEKWCCKGRG 71

DB 564 VQGLSMWCROGOC 576

RESULT 7

AT52_BOVIN STANDARD; PRT; 1205 AA.

AC P79331;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-2 precursor (BC 3.4.24.14) (A disintegrin and metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (Procollagen I/II de amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC I-NP) (Procollagen N-endorpeptidase) (PNPI).

GN ADAMTS2 OR NPI.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RX MEDLINE=97225960; PubMed=9122202;

RA Colige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J., Lapiere C.M.;

RA "cDNA cloning and expression of bovine procollagen I N-proteinase: a new member of the superfamily of zinc-metalloproteinases with binding sites for cells and other matrix components.";

RT Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).

RN [2]

RP PARTIAL SEQUENCE.

RX MEDLINE=95348096; PubMed=7622493;

RA Colige A., Beschin A., Samyn B., Goebels Y., Van Beeumen J., Nusgens B.V., Lapiere C.M.;

RA "Characterization and partial amino acid sequencing of a 107-kDa procollagen I N-proteinase purified by affinity chromatography on immobilized type XIV collagen.";

RT

RL J. Biol. Chem. 270:16724-16730(1995).

CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior to fibril assembly. Does not act on type III collagen. May also play a role in development that is independent of its role in collagen biosynthesis.

CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain alpha-1(I) at Pro-Gln and of alpha-1(II) and alpha-2(I) chains at Ala-Gln.

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to collagen type XIV.

CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Enzymatic activity is detected at high level in all type I collagen-rich tissues such as skin, bones, tendons and aorta and at low level in brain and thymus. The mRNA levels were disproportionately high in heart, liver, retina and muscle.

CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.

CC -!- PTM: The N-terminus is blocked.

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -!- DISEASE: Defects in ADAMTS2 are the cause of dermatosparaxis, a recessively inherited disorder characterized by severe skin fragility and biochemically by the presence in skin of procollagen incompletely processed at the N-terminus.

CC -!- SIMILARITY: Belongs to peptidase family M12B.

CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC -!- SIMILARITY: Contains 1 PLAC domain.

CC -!- SIMILARITY: Contains 4 TSP type-1 domains.

CC -!- CAUTION: Has sometimes been referred to as ADAMTS3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X96389; CAA65253.1; -

DR PIR; T18517; T18517.

DR MEROPS; M12.301; -

DR InterPro; IPR001762; Disintegrin.

DR InterPro; IPR006025; Pept M Zn BS.

DR InterPro; IPR001590; Peptidase_M12B.

DR InterPro; IPR002870; Peptidase_M12B_N.

DR InterPro; IPR000884; TSP1.

DR Pfam; PF01562; Pep_M12B_propep; 1.

DR Pfam; PF01421; Reprolysin; 1.

DR Pfam; PF00093; TSP1; 4.

DR SMART; SMO0209; TSP1; 4.

DR PROSITE; PS50215; ADAM_MEPRO; 1.

DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

DR PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.

DR PROSITE; PS50900; PLAC; 1.

DR PROSITE; PS50092; TSP1; 4.

DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.

DR Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; Repeat; Collagen degradation; Extracellular matrix.

FT SIGNAL 1 28

FT PROPEP 29 253

FT CHAIN 254 1205

FT DOMAIN 254 473

FT DOMAIN 474 554

FT DOMAIN 555 610

FT DOMAIN 612 716

FT DOMAIN 717 845

FT DOMAIN 848 906

FT DOMAIN 908 968

FT DOMAIN 969 1023

FT DOMAIN 1053 1091

FT DOMAIN 31 35

FT DOMAIN 177 190 POLY-GLU.
FT METAL 402 403 ZINC (CATALYTIC) (POTENTIAL).
FT ACT SITE 403 404 BY SIMILARITY.
FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT SITE 685 687 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1092 1092 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1144 1144 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1205 AA; 133887 MW; 755B232A45320371 CRC64;

Query Match 33.8%; Score 135; DB 1; Length 1205;
Best Local Similarity 45.1%; Pred. No. 1.3e-07;
Matches 32; Conservative 5; Mismatches 32; Indels 2; Gaps 2;

QY 1 PGLYSANEQCRVARGPKAVACTFAREHLDMCALCHTDPLQSSCSRLVPLLDGTEC 60
D6 476 PGLHYSMNEQCRFDGLGYMCTAFRT-EDPCQLWC-SHPDNPYFCKTKKGPPLDGTMC 533
QY 61 GVEKWCCKGRC 71
D6 534 APGRHCFKHC 544

RESULT 8
ATS2_HUMAN
ID_ATS2_HUMAN STANDARD; PRT; 1211 AA.
AC O95450;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-2 precursor (BC 3.4.24.14) (A disintegrin and metalloproteinase with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2) (Procollagen I/II DE amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC DE I-NP) (Procollagen N-endopeptidase) (NPPI).
GN ADAMTS2 OR PCINP OR PCPNI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS LENPI AND SPNPI), AND DISEASE.
RC TISSUE=Skin;
RX MEDLINE=99347935; PubMed=10417273;
RA Colige A., Sieron A.L., Li S.-W., Schwarze U., Petty E., Wentelecki W., Wilcox W., Krakow D., Cohn D.H., Reardon W., Byers P.H., Lapiere C.M., Prockop D.J., Nussgens B.V.;
RT "Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene.";
RL Am. J. Hum. Genet. 65:308-317(1999).
CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior to fibril assembly. Does not act on type III collagen. May also play a role in development that is independent of its role in collagen biosynthesis.
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain alpha-1(I) at Pro-Gln and of alpha-1(II) and alpha-2(I) chains at Ala-Gln.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to collagen type XIV (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=lpNPI;
CC IsoId=O95450-1; Sequence=Displayed;
CC Name=spNPI;

CC CC IsoId=O95450-2; Sequence=VSP_005497, VSP_005498;
CC Note=Has no significant N-procollagen peptidase activity;
CC -!- TISSUE SPECIFICITY: Expressed at high level in skin, bone, tendon
CC and aorta and at low levels in thymus and brain.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- DISEASE: Defects in ADAMTS2 are the cause of Ehlers-Danlos
CC syndrome type VIIC (BDS VIIC) [MIM:225410]. BDS VIIC is a
CC recessively inherited connective-tissue disorder characterized
CC clinically by severe skin fragility and joint hypermobility and
CC biochemically by the presence in skin of procollagen incompletely
CC processed at the N-terminus.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- CAUTION: Has sometimes been referred to as ADAMTS3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ003125; CAA05880.1; --
CC MEROPS; M12.301; --
CC Genew; HGNC:218; ADAMTS2.
CC MIM; 604539; --
CC MIM; 225410; --
CC GO; GO:0008133; F:collagenase activity; TAS.
CC GO; GO:0008237; F:metallopeptidase activity; TAS.
CC InterPro; IPR001762; Disintegrin
CC InterPro; IPR006025; Pept_Mz_Bs.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp 1; 4.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS50215; ADAM_MEPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS50300; PLAC; 1.
CC PROSITE; PS50092; TSP1; 4.
CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Collagen degradation; Extracellular matrix;
CC Alternative splicing; Ehlers-Danlos syndrome.
CC SIGNAL 1 29 POTENTIAL.
CC PROPEP 30 253 BY SIMILARITY.
CC CHAIN 254 1211 ADAMTS-2.
CC DOMAIN 254 479 METALLOPROTEASE.
CC DOMAIN 480 560 DISINTEGRIN-LIKE.
CC DOMAIN 561 616 TSP TYPE-1 1.
CC DOMAIN 618 722 CYS-RICH.
CC SITE 691 693 CELL ATTACHMENT SITE (POTENTIAL).
CC DOMAIN 723 851 SPACER.
CC DOMAIN 854 912 TSP TYPE-1 2.
CC DOMAIN 914 971 TSP TYPE-1 3.
CC DOMAIN 975 1029 TSP TYPE-1 4.
CC DOMAIN 1059 1097 PLAC.
CC DOMAIN 40 43 POLY-ALA.
CC DOMAIN 185 188 POLY-GLU.
CC METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT SITE 409 409 BY SIMILARITY.
CC METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).

FT	METAL	418	418	ZINC (CATALYTIC) (BY SIMILARITY).
FT	CARBOHYD	112	112	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	251	251	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	949	949	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	993	993	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1031	1031	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1098	1098	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1145	1145	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	1150	1150	HCFKHCITWLPFDILKRQDSGWA -> FRPGAVAHACPST
FT	VARSPLIC	544	566	LGGGGRWTA (in isoform SpNP1). /FTID=VSP 005497.
FT	VARSPLIC	567	1211	Missing (in isoform SpNP1). /FTID=VSP 005498.
FT	SEQUENCE	1211 AA;	134722 MW;	BECBEF25C3CAD2D CRC64;
Query Match		33.8%;	Score 135;	DB 1; Length 1211;
Best Local Similarity		45.1%;	Pred. No. 1.3e-07;	
Matches 32;	Conservative	5;	Mismatches 32;	Indels 2; Gaps 2;
QY	1 PGLYSANEQCRVFAGKAVACTFAREHLDMCOALSCHTDPLDQSCSRLLVPFLLDGTCC 60			
Db	482 PGLHYSMNEQCRFPFGLYGMCTAFRT-FDKCKQLWC-SHPDNPFCKTKGPFLDGTMC 539			
QY	61 GVEKWCCKSGRC 71			
Db	540 APGKHCFKGHC 550			
RESULT 9				
ID	AT52_MOUSE	STANDARD;	PRT;	1213 AA.
AC	Q8C9W3;			
DT	10-OCT-2003 (Rel. 42, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DE	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase			
DE	with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS 2) (Procollagen I/II			
DE	de amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC			
DE	I-NP) (Procollagen N-endopeptidase) (GNPI).			
GN	ADAMTS2.			
CS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
[[]]				
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Thymus;			
RC	MEDLINE=22354683; PubMed=12466851;			
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,			
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,			
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,			
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,			
RA	Schrim L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,			
RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,			
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,			
RA	Gasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,			
RA	Grimmond S., Gustinchich S., Hirokawa N., Jackson I.J., Jarvis E.D.,			
RA	Kanai A., Kawaji H., Kawasaki Y., Kedzietski R.M., King B.L.,			
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,			
RA	Maglott D.R., Maltas L., Marchionni L., McKenzie L., Miki H.,			
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,			
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,			
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,			
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,			
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,			
RA	Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,			
RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang L., Yang L.,			
RA	Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,			
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Saco K.,			
RA	Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,			
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,			
RA	Miyazaki A., Sakai K., Sagaki D., Shibata K., Shinagawa A.,			

DR PROSITE; PS50900; PLAC; 1.
 DR PROSITE; PS50902; TSP1; 4.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Collagen degradation; Extracellular matrix.
 FT SIGNAL 1 28
 FT PROPEP 29 260
 FT CHAIN 29 1213
 FT DOMAIN 261 471
 FT DOMAIN 480 560
 FT DOMAIN 561 617
 FT DOMAIN 618 722
 FT DOMAIN 723 851
 FT DOMAIN 855 913
 FT DOMAIN 915 975
 FT DOMAIN 976 1030
 FT DOMAIN 1060 1098
 FT SITE 692 694
 FT METAL 409 409
 FT ACT_SITE 410 410
 FT METAL 413 413
 FT METAL 419 419
 FT CARBOHYD 111 111
 FT CARBOHYD 252 252
 FT CARBOHYD 949 949
 FT CARBOHYD 950 950
 FT CARBOHYD 994 994
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1099 1099
 FT CARBOHYD 1147 1147
 FT CARBOHYD 1152 1152
 SQ SEQUENCE 1213 AA; 135298 MW; B27431E00443EDB5 CRC64;

Query Match 33.8%; Score 135; DB 1; Length 1213;
 Best Local Similarity 45.1%; Pred. No. 1.3e-07;
 Matches 32; Conservative 5; Mismatches 32; Indels 2; Gaps 2;

Oy 1 PGLVYSANEQCVAFQPKAVACTFAREHLNMCALSCHTPDLQSSCSRLVPLLDGTEC 60
 Db 483 PGLVYSMEQCFDFGLGYMCTAFRT-PDFCQLWC-SHPDNPYCFKTKGPPLDGTM 540

Oy 61 GVEKWCCKGRC 71
 Db 541 APGKHCFKGC 551

RESULT 10
 ID AT16 HUMAN STANDARD; PRT; 1224 AA.
 AC OSTE7; OSTE2; ...
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-16 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 16) (ADAM-TS 16) (ADAM-TS16).
 GN ADAMTS16 OR KIAA2029.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V., Lopez-Otin C.;
 RA "Cloning, expression analysis, and structural characterization of seven novel human ADAMTS, a family of metalloproteinases with disintegrin and thrombospondin-1 domains.";
 RL Gene 283:49-62(2002).
 RN [2]
 RP SEQUENCE OF 204-1224 FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RA Nagase T., Kikuno R., Ohara O.;

"The nucleotide sequence of a long cDNA clone isolated from human.";
 Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=1;
 CC IsoId=Q8TES7-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8TES7-2; Sequence=VSP_007664, VSP_007665;
 CC -!- TISSUE SPECIFICITY: Expressed in fetal lung and kidney and in adult prostate and ovary.
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix (By similarity).
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 6 TSP type-1 domains.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ315734; CAC86015.1; -;
 CC EMBL; AB095949; BAC23125.1; -;
 CC MEROPS; M12.026; -;
 CC Genew; HGNC:17108; ADAMTS16.
 CC MIM; 607510; -;
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR001818; Pept M12B.
 CC InterPro; IPR006025; Pept M Zn BS.
 CC InterPro; IPR001590; Peptidase_M12B.
 CC InterPro; IPR002870; Peptidase_M12B_N.
 CC InterPro; IPR000884; TSP1.
 CC InterPro; IPR008085; TSP 1.
 CC Pfam; PF01562; Pept M12B_Prophep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC Pfam; PF00090; tse_1; 5.
 CC PRINTS; PR01705; TSP1REPEAT.
 CC SMART; SM00209; TSP1; 6.
 CC PROSITE; PS50215; ADAM_MEPPO; 1.
 CC PROSITE; PS00546; CYSTEINE SWITCH; FALSE_NEG.
 CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
 CC PROSITE; PS50900; PLAC; 1.
 CC PROSITE; PS50092; TSP1; 5.
 CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Alternative splicing.
 FT SIGNAL 1 24
 FT PROPEP 25 279
 FT CHAIN 280 1224
 FT DOMAIN 280 495
 FT DOMAIN 496 585
 FT DOMAIN 586 641
 FT DOMAIN 642 746
 FT DOMAIN 747 873
 FT DOMAIN 874 922
 FT DOMAIN 927 987
 FT DOMAIN 988 1048
 FT DOMAIN 1051 1115
 FT DOMAIN 1127 1181
 FT DOMAIN 1186 1223
 FT SITE 249 249
 FT METAL 433 433

BY SIMILARITY.
 CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.
 CONTAINS 6 TSP TYPE-1 DOMAINS.
 CONTAINS 1 PLAC DOMAIN.
 THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION - THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL ENTITIES REQUIRES A LICENSE AGREEMENT (SEE [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) OR SEND AN EMAIL TO license@isb-sib.ch).

 EMBL; AJ315734; CAC86015.1; -;
 EMBL; AB095949; BAC23125.1; -;
 MEROPS; M12.026; -;
 Genew; HGNC:17108; ADAMTS16.
 MIM; 607510; -;
 InterPro; IPR001762; Disintegrin.
 InterPro; IPR001818; Pept M12B.
 InterPro; IPR006025; Pept M Zn BS.
 InterPro; IPR001590; Peptidase_M12B.
 InterPro; IPR002870; Peptidase_M12B_N.
 InterPro; IPR000884; TSP1.
 InterPro; IPR008085; TSP 1.
 Pfam; PF01562; Pept M12B_Prophep; 1.
 Pfam; PF01421; Reprolysin; 1.
 Pfam; PF00090; tse_1; 5.
 PRINTS; PR01705; TSP1REPEAT.
 SMART; SM00209; TSP1; 6.
 PROSITE; PS50215; ADAM_MEPPO; 1.
 PROSITE; PS00546; CYSTEINE SWITCH; FALSE_NEG.
 PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.
 PROSITE; PS50900; PLAC; 1.
 PROSITE; PS50092; TSP1; 5.
 PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 Repeat; Extracellular matrix; Alternative splicing.
 SIGNAL 1 24
 PROPEP 25 279
 CHAIN 280 1224
 DOMAIN 280 495
 DOMAIN 496 585
 DOMAIN 586 641
 DOMAIN 642 746
 DOMAIN 747 873
 DOMAIN 874 922
 DOMAIN 927 987
 DOMAIN 988 1048
 DOMAIN 1051 1115
 DOMAIN 1127 1181
 DOMAIN 1186 1223
 SITE 249 249
 METAL 433 433

```
FT ACT SITE 434 434 BY SIMILARITY. (BY SIMILARITY).
FT METAL 437 437 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 443 443 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 741 741 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 780 780 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 835 835 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 905 905 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 935 935 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1063 1072 CSVTCTGCTG -> VGLVSRERG (in isoform 2).
FT VARSPLIC 1073 1224 /FTID=VSP_007664.
FT VARSPLIC 1224 Missing (in isoform 2).
FT CONFLICT 284 284 /FTID=VSP_007665.
FT CONFLICT 1224 AA; 136186 NW; 572E72129E1401F8 CRC64;
SQ SEQUENCE 1224 AA; 136186 NW; 572E72129E1401F8 CRC64;

Query Match 33.8%; Score 135; DB 1; Length 1224;
Best Local Similarity 39.4%; Pred No. 1.3e-07;
Matches 28; Conservative 8; Mismatches 31; Indels 4; Gaps 2;

QY 1 PGLYSSANQCRVAFQPKAVACTFAREHLDNCQALSCHTDPLDQSSCSRLVPLLDGTGC 60
Dd 508 PGLYLDANTQCKWQFGKAKLMLDFKK-DICKALWCHR---IGRKCTKMPAEGTIC 563

QY 61 GVEKWCCKGRC 71
Dd 564 GHDWMCRCGQC 574

RESULT 11
AT14 HUMAN
ID AT14 HUMAN STANDARD; PRT; 1223 AA.
AC Q8WXS8; Q8TE55; Q8TEY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-14 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Brederlow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
RT metalloproteinase family.";
RL Biochim. Biophys. Acta 1522:221-225(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Fetal lung;
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
RN [3]
RP SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND
RP ALTERNATIVE PROMOTER USAGE.
RX MEDLINE=21839041; PubMed=11741898;
RA Collige A., Vandenbergh I., Thiry M., Lambert C.A., Van Beumen J.,
RA Li S.-W., Prockop D.J., Lapiere C.M., Nuygens B.V.;
RT "Cloning and characterization of ADAMTS-14, a novel
RT high homology with ADAMTS-2 and ADAMTS-3.";
RL J. Biol. Chem. 277:5756-5766(2002).
CC -!- FUNCTION: Has a aminoprocollagen type I activity processing
CC activity in the absence of ADAMTS2. Seems to be synthesized as a
CC latent enzyme that requires activation to display aminoprocollagen
```


FT DOMAIN 730 846 SPACER.
FT DOMAIN 847 907 TSP TYPE-1 2.
FT DOMAIN 908 967 TSP TYPE-1 3.
FT DOMAIN 968 1022 TSP TYPE-1 4.
FT DOMAIN 1059 1097 PLAC.
FT DOMAIN 875 878 POLY-ARG.
FT DOMAIN 1100 1221 PRO-RICH.
FT ACT_SITE 399 399 BY SIMILARITY.
FT METAL 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 941 941 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 68 Missing (in isoform B and isoform C).
FT VARSPLIC 1 68 /FTID=VSP_006958.
FT VARSPLIC 368 368 /FTID=VSP_005501.
FT CONFLICT 868 868 Q -> R (IN REF. 2).
FT CONFLICT 884 884 Q -> H (IN REF. 2).
FT CONFLICT 901 901 C -> S (IN REF. 2).
FT CONFLICT 923 923 C -> Y (IN REF. 2).
FT CONFLICT 1024 1024 N -> S (IN REF. 2).
SQ SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F170F CRC64;
Query Match 33.6%; Score 134; DB 1; Length 1223;
Best Local Similarity 42.3%; Pred. No. 1.7e-07;
Matches 30; Conservative 6; Mismatches 33; Indels 2; Gaps 2;
Qy 1 PGLYSANEQCRVAFGPKAVACTFAREHLMCOALSCHTDPLQSSCSRLVPLLDGTEC 60
Db 472 PGINYSMDQCRFFDGSYGQTC-LARFTEPCQLWC-SHPDNPYFCKTKKGGPLDGTGEC 529
Qy 61 GVEKWCSGKRC 71
Db 530 APGKWCFKGHC 540
RESULT 12
AT20 HUMAN
ID AT20 HUMAN STANDARD; PRT; 1911 AA.
AC P59510;
DT 10-OCT-2003 (Rel. 42, Created)
DR 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
DE TS20)
GN ADAMTS20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wight T.N., Leduc R., Apte S.S.;
RA "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RA subfamily related to Caenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513(2003).
RL [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Liver;
RX MEDLINE=22566039; PubMed=12562771;
RA Llamazares M., Cal S., Quesada V., Lopez-Otin C.;
RA "Identification and characterization of ADAMTS-20 defines a novel
RA subfamily of metalloproteinases-disintegrins with multiple
RA thrombospondin-1 repeats and a unique GON domain.";
RL J. Biol. Chem. 278:13382-13389(2003).
CC -!- FUNCTION: May play a role in tissue-remodeling process occurring
CC in both normal and pathological conditions.

CC -!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59510-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59510-2; Sequence=VSP_007106, VSP_007107, VSP_007108;
CC -!- TISSUE SPECIFICITY: Very sparingly expressed, although is detected
CC at low levels in testis, prostate, ovary, heart, placenta, lung
CC and pancreas. Overexpressed in several brain, colon and breast
CC carcinomas.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF488804; AAO15766.1; -;
CC EMBL; AJ515153; CAD56159.3; -;
CC EMBL; AJ515154; CAD56160.2; -;
CC Genew; HGNC:17178; ADAMTS20.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSPI.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.
CC Pfam; PF00090; tsp_1; 11.
CC SMART; SM00209; TSPI; 12.
CC PROSITE; PS0215; ADAM_MEPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00092; TSPI; 12.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Repeat; Extracellular matrix; Alternative splicing.
CC SIGNAL 1 21 POTENTIAL.
CC PROPEP 22 253 BY SIMILARITY.
CC CHAIN 254 1911 ADAMTS-20.
CC DOMAIN 254 467 METALLOPROTEASE.
CC DOMAIN 468 555 DISINTEGRIN-LIKE.
CC DOMAIN 556 611 TSP TYPE-1 1.
CC DOMAIN 612 723 CYS-RICH.
CC DOMAIN 724 846 SPACER.
CC DOMAIN 847 905 TSP TYPE-1 2.
CC DOMAIN 906 962 TSP TYPE-1 3.
CC DOMAIN 967 1024 TSP TYPE-1 4.
CC DOMAIN 1025 1074 TSP TYPE-1 5.
CC DOMAIN 1077 1136 TSP TYPE-1 6.
CC DOMAIN 1153 1207 TSP TYPE-1 7.
CC DOMAIN 1208 1265 TSP TYPE-1 8.
CC DOMAIN 1305 1357 TSP TYPE-1 9.
CC DOMAIN 1359 1417 TSP TYPE-1 10.
CC DOMAIN 1418 1476 TSP TYPE-1 11.
CC DOMAIN 1477 1532 TSP TYPE-1 12.
CC DOMAIN 1536 1589 TSP TYPE-1 13.
CC DOMAIN 1590 1653 TSP TYPE-1 14.
CC DOMAIN 1655 1711 TSP TYPE-1 15.
CC DOMAIN 1712 1911 GON.
CC METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).
CC ACT_SITE 404 404 BY SIMILARITY.

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

CC -----

CC EMBL; AF140675; AAD56358.1; -

CC RSP; P15167; IATL.

CC MEROPS; M12.231; -

CC Genew; HGNC:223; ADAMTS7.

CC MIM; 605009; -

CC GO; GO:0008237; P:metallopeptidase activity; TAS.

CC InterPro; IPR001762; Disintegrin.

CC InterPro; IPR001818; Pept_M10A_M12B.

CC InterPro; IPR006025; Pept_M_Zn_BS.

CC InterPro; IPR001590; Peptidase_M12B.

CC InterPro; IPR002870; Peptidase_M12B_N.

CC InterPro; IPR000884; TSP1.

CC InterPro; IPR008085; TSP_1.

CC Pfam; PF01562; Pep_M12B_propep; 1.

CC Pfam; PF01421; Reprolysin; 1.

CC Pfam; PF00090; tsp_1; 1.

CC PRINTS; PR01705; TSP1REPEAT.

CC SMART; SM00209; TSP1; 2.

CC PROSITE; PS00142; ZINC_PROTEASE; 1.

CC PROSITE; PS0215; ADAM_MEPRO; 1.

CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.

CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.

CC PROSITE; PS00092; TSP1; 1.

CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;

KW Repeat; Extracellular matrix.

FT SIGNAL 1 27 POTENTIAL.

FT PROPEP 28 232 BY SIMILARITY.

FT CHAIN 233 997 ADAMTS-7.

FT DOMAIN 233 461 METALLOPROTEASE.

FT DOMAIN 462 537 DISINTEGRIN-LIKE.

FT DOMAIN 538 593 TSP TYPE-1 1.

FT DOMAIN 595 697 CYS-RICH.

FT DOMAIN 698 914 SPACER.

FT DOMAIN 915 974 TSP TYPE-1 2.

FT SITE 204 204 CYSTEINE SWITCH (POTENTIAL).

FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).

FT ACT_SITE 389 389 BY SIMILARITY.

FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 398 398 ZINC (CATALYTIC) (BY SIMILARITY).

FT CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 693 693 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 778 778 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 997 AA; 109694 MW; 6587044ED02FC104 CRC64;

Query Match 31.2%; Score 124.5; DB 1; Length 997;

Best Local Similarity 34.7%; Pred. No. 1.7e-06;

Matches 25; Conservative 10; Mismatches 28; Indels 9; Gaps 3;

QY 1 PGLYSANECQVAFGPKVACTFAREHLD-MCQALSCHTDPLDQSSCSRLVPLLDGTE 59

DB 464 PGLYDVSHQCRLOYGAYSAPC-----EDMDNVCHTLWSVG-----TTCHSKLDAVDGTR 515

QY 60 CGVEKWCCKGRC 71

DB 516 CGENKWCCKSGEC 527

RESULT 15

AT15_MOUSE STANDARD; PRT; 562 AA.

AC P59384;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15) (Fragment).

DE

FT DOMAIN 911 967 TSP TYPE-1 3.

FT SITE 194 198 POLY-ARG.

FT METAL 205 205 CYSTEINE_SWITCH (POTENTIAL).

FT ACT_SITE 401 401 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 402 402 BY SIMILARITY.

FT METAL 405 405 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 411 411 ZINC (CATALYTIC) (BY SIMILARITY).

FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 720 720 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 764 764 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 782 782 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CONFLICT 21 21 I -> V (IN REF. 2).

FT CONFLICT 26 31 KFRSSQ -> RSRGSL (IN REF. 2).

FT CONFLICT 49 49 V -> A (IN REF. 2).

FT CONFLICT 72 72 R -> P (IN REF. 2).

FT CONFLICT 79 79 L -> TR (IN REF. 2).

FT CONFLICT 243 249 R -> G (IN REF. 2).

FT CONFLICT 262 265 TMLV -> G (IN REF. 2).

FT CONFLICT 607 607 S -> F (IN REF. 2).

FT CONFLICT 936 936 L -> V (IN REF. 2).

FT CONFLICT 962 962 I -> T (IN REF. 2).

SQ SEQUENCE 967 AA; 105705 MW; F93C86AF6DCDB4CF CRC64;

Query Match 31.2%; Score 124.5; DB 1; Length 967;

Best Local Similarity 36.6%; Pred. No. 1.6e-06;

Matches 26; Conservative 4; Mismatches 38; Indels 3; Gaps 1;

QY 1 PGLYSANECQVAFGPKVACTFAREHLD-MCQALSCHTDPLDQSSCSRLVPLLDGTEC 60

DB 478 PGLYDANRQCQFTGFEESTHCPDA---ASTCTLWCTGTSGLLVCQTKHPFWADGTSC 534

QY 61 GVEKWCCKGRC 71

DB 535 GEGKWCCKGRC 545

RESULT 14

AT17_HUMAN STANDARD; PRT; 997 AA.

AC Q9UKP4;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-7 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 7) (ADAM-TS 7) (ADAM-TS7).

GN ADAMTS7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99395124; PubMed=10464289;

RA Hurskainen T.L., Hirohata S., Seidin M.F., Apte S.S.;

RT "ADAM-TS5, ADAM-TS6, and ADAM-TS7, novel members of a new family of zinc metalloproteases.";

RL J. Biol. Chem. 274:25555-25563 (1999).

CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -!- TISSUE SPECIFICITY: Expressed in heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas.

CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.

CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -!- SIMILARITY: Belongs to peptidase family M12B.

CC -!- SIMILARITY: Contains 1 disintegrin-like domain.

CC -!- SIMILARITY: Contains 2 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

GN ADAMTS15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EVB/N; TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC043308; AAH43308.1; -.
DR MGD; MGI:2449569; Adamts15.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF00090; tsp_1; 3.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MEPPO; PARTIAL.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; PARTIAL.
KW Hydrolase; Metalloprotease; Zinc; Glycoprotein; Repeat;
KW Extracellular matrix.
FT NON TER 1
FT DOMAIN <1 39 METALLOPROTEASE.
FT DOMAIN 40 127 DISINTEGRIN-LIKE.

FT DOMAIN 128 183 TSP TYPE-1 1.
FT DOMAIN 184 312 CYS-RICH.
FT DOMAIN 311 450 SPACER.
FT DOMAIN 451 507 TSP TYPE-1 2.
FT DOMAIN 508 561 TSP TYPE-1 3.
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 562 AA; 60963 MW; 48009AD337D44E4A CRC64;

Query Match 31.1%; Score 124; DB 1; Length 562;
Best Local Similarity 36.8%; Pred. No. 1.1e-06;
Matches 26; Conservative 7; Mismatches 34; Indels 4; Gaps 1;

QY 1 PGLYSGANQQRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTSC 60
Db 50 PGTYSYLSQOCELAFVGSKPCPY----MQYCTKLWCTGKAKGQWVCQTRHPPWADGTSC 105

QY 61 GVEKWCQKRC 71
Db 106 GEGKFCCLKGAC 116

Search completed: March 13, 2004, 07:39:54
Job time : 2.47917 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 7.02604 Seconds
(without alignments)
3188.394 Million cell updates/sec

Title: US-09-836-712-2_COPY_324_394

Perfect score: 399

Sequence: 1 POLYSSANEQCRVAFGPKAV.....VPLLDGTGCGVKWCSKGRG 71

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL 25:
1: sp_archaea:
2: sp_bacteria:
3: sp_fungi:
4: sp_human:
5: sp_invertebrate:
6: sp_mammal:
7: sp_mhc:
8: sp_organelle:
9: sp_phase:
10: sp_plant:
11: sp_rodent:
12: sp_virus:
13: sp_vertebrate:
14: sp_unclassified:
15: sp_rvirus:
16: sp_bacteriap:
17: sp_archaeap:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	399	100.0	1427	Q96L37	Q96L37 homo sapien
2	242	60.7	364	Q9UCQ1	Q9UCQ1 homo sapien
3	141.5	35.5	1009	Q8BK11	Q8BK11 mus musculus
4	141.5	35.5	1600	Q811B3	Q811B3 mus musculus
5	137	34.3	900	Q8K206	Q8K206 mus musculus
6	134	33.6	1092	Q8BK11	Q8BK11 mus musculus
7	126	31.6	790	Q8T458	Q8T458 drosophila
8	126	31.6	790	Q8MM11	Q8MM11 drosophila
9	126	31.6	1091	Q9W126	Q9W126 drosophila
10	124.5	31.2	509	Q20930	Q20930 caenorhabdi
11	124	31.1	269	Q9GL54	Q9GL54 cryctolagus
12	122	30.6	623	Q8BGP4	Q8BGP4 mus musculus
13	121	30.3	187	Q95N23	Q95N23 equus cabal
14	119.5	29.9	967	Q8NE26	Q8NE26 homo sapien
15	118.5	29.7	759	Q8H2M8	Q8H2M8 equus cabal
16	117	29.3	975	Q7Z291	Q7Z291 caenorhabdi

17	117	29.3	1020	5	Q8IU50	Q8IU50 caenorhabdi
18	117	29.3	1020	5	Q19204	Q19204 caenorhabdi
19	112.5	28.2	769	5	Q8MRL5	Q8MRL5 drosophila
20	112.5	28.2	1059	5	Q9W493	Q9W493 drosophila
21	111	27.8	2165	5	Q19791	Q19791 caenorhabdi
22	110	27.6	1688	5	Q8SXB0	Q8SXB0 drosophila
23	109.5	27.4	833	11	Q8K384	Q8K384 mus musculus
24	109.5	27.4	845	11	Q8BNJ2	Q8BNJ2 mus musculus
25	107.5	26.9	192	6	Q95N24	Q95N24 equus cabal
26	107.5	26.9	203	11	Q8CH80	Q8CH80 cavia porce
27	107.5	26.9	839	6	Q7YS95	Q7YS95 bos taurus
28	101.5	25.4	1461	5	Q8MYA8	Q8MYA8 caenorhabdi
29	97	24.3	108	6	Q9GLK6	Q9GLK6 cryctolagus
30	96.5	24.2	1070	11	Q8CG28	Q8CG28 mus musculus
31	78	19.5	610	13	Q9DGB9	Q9DGB9 crotalus at
32	75	18.8	610	13	Q8AW15	Q8AW15 agkistrodon
33	74.5	18.7	872	5	Q22580	Q22580 caenorhabdi
34	73.5	18.4	247	17	Q8PY45	Q8PY45 methanosarc
35	72.5	18.2	730	11	Q80Y08	Q80Y08 mus musculus
36	70.5	17.7	112	2	Q52067	Q52067 pantoea agg
37	68.5	17.2	754	11	Q8CGQ1	Q8CGQ1 mus musculus
38	68	17.0	620	13	Q42138	Q42138 agkistrodon
39	67.5	16.9	609	13	Q8UGN1	Q8UGN1 naja mossam
40	67	16.8	313	15	O56863	O56863 feline foam
41	67	16.8	387	15	O93036	O93036 feline foam
42	67	16.8	488	5	Q9GP14	Q9GP14 ixodes rici
43	66	16.5	209	4	Q86VD6	Q86VD6 homo sapien
44	66	16.5	455	16	Q92LU7	Q92LU7 rhizobium m
45	66	16.5	512	2	O85725	O85725 streptomyce

ALIGNMENTS

RESULT 1

Q96L37 PRELIMINARY; PRT; 1427 AA.

AC Q96L37 (TREMEL. 19, Created)
DT 01-DEC-2001 (TREMEL. 19, Last sequence update)
DT 01-DEC-2001 (TREMEL. 25, Last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E.,
RA Fujikawa K.;

RT "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a
RT Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura.";
RL J. Biol. Chem. 276:41059-41063(2001).

DR EMBL; AY055376; AAL17652.1;
DR MEROPS; M12.241;
DR Genew; HGNC:1366; ADAMTS13.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSPL.
DR Pfam; PF01421; Reprolysin; 1.
DR SMART; SM00209; TSP1; 7.
DR PROSITE; PS50215; ADAM_MEPPO; 1.
DR PROSITE; PS50092; TSPL; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Signal.
FT SIGNAL 1 33 POTENTIAL.

FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; E81BC3AABC1A4442 CRC64;
Query Match 100.0%; Score 399; DB 4; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1e-43;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PGLYTSANEQCRVAFQPKAVACTFAREHLMDCQALSCHTDPLDQSSCSRLVPLLDGTGTC 60
DB 301 PGLYTSANEQCRVAFQPKAVACTFAREHLMDCQALSCHTDPLDQSSCSRLVPLLDGTGTC 360
QY 61 GVEKWCCKGRC 71
DB 361 GVEKWCCKGRC 371
RESULT 2
ID Q9UQ01 PRELIMINARY; PRT; 364 AA.
AC Q9UQ01;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN C9ORF8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Young J.M., Woodward K.J., Aziz S., Burley M., Kwiatkowski D.J.,
Povey S.;
RT "Cloning of a sugar transporter gene, a G-beta subunit like gene and
three novel genes in human chromosome 9q34.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011374; CAB66157.1;
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP_1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1.1.
DR PROSITE; PS50092; TSP1.1.
DR Hypothetical protein.
SQ SEQUENCE 364 AA; 39864 MW; DA42FC5F5345F3A0 CRC64;
Query Match 60.7%; Score 242; DB 4; Length 364;
Best Local Similarity 97.7%; Pred. No. 1.5e-23;
Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTGTCGVEKWCCKGRC 71
DB 1 LDMCQALSCHTDPLDQSSCSRLVPLLDGTGTCGVEKWCCKGRC 43
RESULT 3
ID Q8BKV1 PRELIMINARY; PRT; 1009 AA.
AC Q8BKV1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ADAMTS-12 precursor (Fragment).
GN AI605170.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK048612; BAC33391.1; --
DR PIR; PT0546; PT0698.
DR MGD; MGI:2146046; AI605170.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR00885; TSP_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 4.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1.1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1.3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1009 1009
SQ SEQUENCE 1009 AA; 113352 MW; 7B63218CFF0FDB1 CRC64;
Query Match 35.5%; Score 141.5; DB 11; Length 1009;
Best Local Similarity 38.0%; Pred. No. 8.3e-10;
Matches 27; Conservative 11; Mismatches 26; Indels 7; Gaps 2;
QY 1 PGLYTSANEQCRVAFQPKAVACTFAREHLMDCQALSCHTDPLDQSSCSRLVPLLDGTGTC 60
DB 472 PGVIVDVHQCQLQYGFNA---TFCQEVENVCTLWCSSV---KGFCSKLDAAADGTGTC 524
QY 61 GVEKWCCKGRC 71
DB 525 GEKKWCCKGRC 535
RESULT 4
ID Q811B3 PRELIMINARY; PRT; 1600 AA.
AC Q811B3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloprotease disintegrin 12 protein.
GN ADAMTS-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Cal S., Lopez-Otin C.;
RT "Mouse ADAMTS-12.";
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ537452; CAD60967.1; --
DR PIR; PT0546; PT0698.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:000508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR00884; TSP1.
DR InterPro; IPR00885; TSP_1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 7.
DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00209; TSP1; 8.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Integrin; Protease.
SQ SEQUENCE 1600 AA; 177791 MW; 083089D356E38C5D CRC64;

Query Match 35.5%; Score 141.5; DB 11; Length 1600;
Best Local Similarity 38.0%; Pred. No. 1.3e-09;
Matches 27; Conservative 11; Mismatches 26; Indels 7; Gaps 2;

Qy 1 POLYISANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTTC 60
Db 472 PGVIYDVHQQQLQYGFNA---TFCQEVENVQTLWGSV---KGFCSRKLDAADGTRC 524
Qy 61 GVEKWCCKGRC 71
Db 525 GEKKWCMAGKC 535

RESULT 5
Q8K206 PRELIMINARY; PRT; 900 AA.
ID Q8K206;
AC Q8K206;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ADAMTS16.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034739; AH34739.1; -.
DR MGI; MGI:2429637; Adamts16.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP.1.
DR Pfam; PF01562; Pep_M12B_prosep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Hypothetical protein.
SQ SEQUENCE 900 AA; 100679 MW; 590BE2B0E73740FF CRC64;

Query Match 34.3%; Score 137; DB 11; Length 900;
Best Local Similarity 39.7%; Pred. No. 2.9e-09;
Matches 29; Conservative 8; Mismatches 28; Indels 8; Gaps 3;

Qy 1 POLYISANEQCRVAFGPKAVAC--TFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGT 58
Db 506 PGQLYDANTCKQKQFGKAKLMLDFRK---DICKALWCHR---IGRCETKFMFAEAGT 559
Qy 59 ECGVEKWCCKGRC 71
Db 560 LCQQDMWCRGGQC 572

RESULT 6
Q8BKAI PRELIMINARY; PRT; 1092 AA.
ID Q8BKAI
AC Q8BKAI;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical metalloprotease.
GN E130314N14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Eye;
RX MEDLINE=22354693; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK053851; BAC35556.1; -.
DR MGI; MGI:2442600; E130314N14RIK.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP.1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 4.
KW Hypothetical protein.
SQ SEQUENCE 1092 AA; 120324 MW; B5F03956553536AE CRC64;

Query Match 33.8%; Score 134; DB 11; Length 1092;
Best Local Similarity 36.6%; Pred. No. 8.9e-09;
Matches 26; Conservative 11; Mismatches 30; Indels 4; Gaps 2;

Qy 1 POLYISANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTTC 60
Db 339 PGQYDADQCKQKQFGAKKCLSLG--YMKOICKSLWCHR---VGHRCETKFMFAEAGTAC 394
Qy 61 GVEKWCCKGRC 71
Db 395 GLSMWCRGGQC 405

RESULT 7
Q8T458 PRELIMINARY; PRT; 790 AA.
ID Q8T458;
AC Q8T458;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT15733P.
GN CG3622 OR BCDNA:GM08694.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champs M., Chavez C., Dorsett V., Dresnek D., Farfan D., Friese E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleeb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089340; AAL90078.1; -.
DR FlyBase; FBgn0034778; CG3622.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.

DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR006025; Pept_M_Zn_BS.
DR InterPro: IPR000884; TSPI.
DR Pfam: PF01421; Reprolysin; 1.
DR PROSITE: PS50215; ADAM_MBP; 1.
DR PROSITE: PS50092; TSPI; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 790 AA; 89226 MW; DF8ADB792D78E794 CRC64;

Query Match 31.6%; Score 126; DB 5; Length 790;
Best Local Similarity 35.2%; Pred. No. 7.2e-08;
Matches 25; Conservative 13; Mismatches 27; Indels 6; Gaps 2;

QY 1 PGLYSANEQCRVAFGPKVACTFAREHLDVQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 215 PGERFNANQCMLEFRFGKFNQAS-TQSKMEICRDLHCRQDGLPWTs-----HPALEGTEC 268

QY 61 GVEKWCCKGRC 71
Db 269 GPNMWCRCGGTC 279

RESULT 8
Q8MME1 PRELIMINARY; PRT; 790 AA.
AC Q8MME1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG3622-PA.
GN CG3622 OR BCDNA:GM08694.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattai B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Amanatides P.G., Brandon R.C., Rogers Y., An H., Baldwin D.,
RA Banson J., Beeson K.Y., Busam D.A., Carlson J.W., Center A.,
RA Champe M., Davenport L.B., Dietz S.M., Dodson K., Dorsett V.,
RA Doup L.E., Doyle C., Dresnek D., Farfan D., Ferreira S., Frise E.,
RA Galle R.F., Garg N.S., George R.A., Gonzalez M., Houck J.,
RA Hoskins R.A., Hostin D., Howland T.J., Ibegwan C., Jalali M.,
RA Kruse D., Li P., Mattai B., Moshrefi A., McIntosh T.C., Moy M.,
RA Murphy B., Nelson C., Nelson K.A., Nuccio J., Pacleb J., Paragas V.,
RA Park S., Patel S., Pfeiffer F., Phouanavong S., Pittman G.S.,
RA Puri V., Richards S., Scheeler F., Stapleton M., Strong R.,
RA Swirskas R., Tector C., Tyler D., Williams S.M., Zaveri J.S.,
RA Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome";
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celnik S.E.,
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smatnik F., Rubin G.M.,
RA Whitfield E.J., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome";
Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AS003458; AM71125.1; -;
DR FlyBase; FBgn0034778; CG3622.
DR GO: GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO: GO:0008270; F:zinc ion binding; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001590; Peptidase_M12B.
DR InterPro: IPR006025; Pept_M_Zn_BS.
DR InterPro: IPR000884; TSPI.
DR Pfam: PF01421; Reprolysin; 1.
DR PROSITE: PS50215; ADAM_MBP; 1.
DR PROSITE: PS50092; TSPI; 1.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 790 AA; 89319 MW; 48FC0539882A9C8A CRC64;

Query Match 31.6%; Score 126; DB 5; Length 790;
Best Local Similarity 35.2%; Pred. No. 7.2e-08;
Matches 25; Conservative 13; Mismatches 27; Indels 6; Gaps 2;

QY 1 PGLYSANEQCRVAFGPKVACTFAREHLDVQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 215 PGERFNANQCMLEFRFGKFNQAS-TQSKMEICRDLHCRQDGLPWTs-----HPALEGTEC 268

QY 61 GVEKWCCKGRC 71
Db 269 GPNMWCRCGGTC 279

RESULT 9
Q9W1Z6 PRELIMINARY; PRT; 1091 AA.
ID Q9W1Z6
AC Q9W1Z6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CG3622 protein.
GN CG3622 OR BCDNA:GM08694.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

[illegible]

```
RT chondrocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317415; AAG33062.1; -
DR HSSP; O9PW35; 1BUD.
DR MEROPS; M12.225; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP.1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1.1.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS0092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

Query Match 31.1%; Score 124; DB 6; Length 269;
Best Local Similarity 33.8%; Pred. No. 4.3e-08;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 110 PGQTYDATQCCLNTFGPEYTVG-----PGMDVCAFLWCAVVRQGMVCLTKLPVAVGTGTPC 165
QY 61 GVEKWCCKGRC 71
Db 166 GKGRICLQKGC 176

RESULT 12
Q8BGP4 PRELIMINARY; PRT; 623 AA.
AC Q8BGP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease (Fragment).
GN ADAMTS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK04746; BAC32062.1; -
DR EMBL; AK082475; BAC38503.1; -.
DR MGD; MGI:1346321; Adamts5.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR002088; PPTA-.
DR InterPro; IPR008084; TSP1.
DR Pfam; PF01421; Reprolysin; 1.

RT chondrocytes."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF317415; AAG33062.1; -
DR HSSP; O9PW35; 1BUD.
DR MEROPS; M12.225; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP.1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1.1.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS0092; TSP1; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

Query Match 31.1%; Score 124; DB 6; Length 269;
Best Local Similarity 33.8%; Pred. No. 4.3e-08;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 110 PGQTYDATQCCLNTFGPEYTVG-----PGMDVCAFLWCAVVRQGMVCLTKLPVAVGTGTPC 165
QY 61 GVEKWCCKGRC 71
Db 166 GKGRICLQKGC 176

RESULT 12
Q8BGP4 PRELIMINARY; PRT; 623 AA.
AC Q8BGP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease (Fragment).
GN ADAMTS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK04746; BAC32062.1; -
DR EMBL; AK082475; BAC38503.1; -.
DR MGD; MGI:1346321; Adamts5.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR002088; PPTA-.
DR InterPro; IPR008084; TSP1.
DR Pfam; PF01421; Reprolysin; 1.

chondrocytes."
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AF317415; AAG33062.1; -
HSSP; O9PW35; 1BUD.
MEROPS; M12.225; -.
GO; GO:0004222; F:metalloendopeptidase activity; IEA.
GO; GO:0008270; F:zinc ion binding; IEA.
GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
InterPro; IPR001590; Peptidase M12B.
InterPro; IPR006025; Pept_M_Zn_BS.
InterPro; IPR000884; TSP1.
InterPro; IPR008085; TSP.1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; tsp.1; 1.
PRINTS; PR01705; TSP1REPEAT.
SMART; SM00209; TSP1.1.
PROSITE; PS0215; ADAM_MEPRO; 1.
PROSITE; PS0092; TSP1; 1.
PROSITE; PS00142; ZINC_PROTEASE; 1.
NON_TER 1
FT NON_TER 269
SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;

Query Match 31.1%; Score 124; DB 6; Length 269;
Best Local Similarity 33.8%; Pred. No. 4.3e-08;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 110 PGQTYDATQCCLNTFGPEYTVG-----PGMDVCAFLWCAVVRQGMVCLTKLPVAVGTGTPC 165
QY 61 GVEKWCCKGRC 71
Db 166 GKGRICLQKGC 176

RESULT 12
Q8BGP4 PRELIMINARY; PRT; 623 AA.
AC Q8BGP4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE A disintegrin-like and metalloprotease (Fragment).
GN ADAMTS5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Retina;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK04746; BAC32062.1; -
DR EMBL; AK082475; BAC38503.1; -.
DR MGD; MGI:1346321; Adamts5.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR002088; PPTA-.
DR InterPro; IPR008084; TSP1.
DR Pfam; PF01421; Reprolysin; 1.
```

```
DR Pfam; PF00090; tsp.1; 2.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00904; PPTA; 1.
DR PROSITE; PS0092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
SQ SEQUENCE 623 AA; 68681 MW; 80CE9BCD0AC37502 CRC64;

Query Match 30.6%; Score 122; DB 11; Length 623;
Best Local Similarity 33.8%; Pred. No. 1.9e-07;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 180 PGQTYDATQCCLNTFGPEYSVC-----PGMDVCAFLWCAVVRQGMVCLTKLPVAVGTGTPC 235
QY 61 GVEKWCCKGRC 71
Db 236 GKGRVCLQKGC 246

RESULT 13
Q95N23 PRELIMINARY; PRT; 187 AA.
AC Q95N23;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aggrecanase-2 (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Flannery C.R., Little C.B.;
RT "Expression and activity of equine aggrecanases."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF368322; AAK53426.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 1.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS0092; TSP1; 1.
FT NON_TER 1
FT NON_TER 187
SQ SEQUENCE 187 AA; 20477 MW; DB10785095372E8C CRC64;

Query Match 30.3%; Score 121; DB 6; Length 187;
Best Local Similarity 33.8%; Pred. No. 7.3e-08;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLMQCALSCHTDPLDQSSCSRLVPLLDGTEC 60
Db 73 PGQTYDASQCCLNTFGPEYSVC-----PGMDVCAFLWCAVVRQGMVCLTKLPVAVGTGTPC 128
QY 61 GVEKWCCKGRC 71
Db 129 GKGRICLHKGK 139

RESULT 14
Q8NE26 PRELIMINARY; PRT; 967 AA.
AC Q8NE26;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

```
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC036515; AAH36515.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001128; Cytochrome_P450.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01562; Pcp_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp 1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR PROSITE; PS00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MBP; 1.
DR PROSITE; PS00886; CYTOCHROME_P450; 1.
DR PROSITE; PS0092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hypothetical protein.
SQ SEQUENCE 967 AA; 105387 MW; FF1D399674201C3D CRC64;

Query Match 29.9%; Score 119.5; DB 4; Length 967;
Best Local Similarity 36.6%; Pred. No. 6.5e-07;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYSSANEQCRVAFGPKAVACTFAREHLDLQSCSRLVPLLDGTGC 60
DB 478 PGTSDANRQCQFTFGESKXCPDA---ASTCTLTWCTGSGGLVLCQTKHFFWADGTSC 534

QY 61 GVEKWCCKGRC 71
DB 535 GEGKWCINGKC 545

RESULT 15
Q8H2M8 PRELIMINARY; PRT; 759 AA.
AC Q8H2M8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloproteinase (Fragment).
GN ADAMTS-1.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Boerboom D., Russell D.L., Richards J.S., Sirois J.;
RT "Coordinated Regulation of Transcripts Encoding ADAMTS-1 and
RT Progesterone Receptor in Equine Preovulatory Follicles.";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF541975; AAN17331.1; -.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR006586; ADAM_cysteine.
DR InterPro; IPR001590; Peptidase_M12B.
```

```
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp 1; 3.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS0215; ADAM_MBP; 1.
DR PROSITE; PS0092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
FT NON_TER 1
SQ SEQUENCE 759 AA; 83410 MW; C18E4C048918C8E9 CRC64;

Query Match 29.7%; Score 118.5; DB 6; Length 759;
Best Local Similarity 35.2%; Pred. No. 6.8e-07;
Matches 25; Conservative 5; Mismatches 38; Indels 3; Gaps 1;

QY 1 PGLYSSANEQCRVAFGPKAVACTFAREHLDLQSCSRLVPLLDGTGC 60
DB 270 PGTLYDANRQCQFTFGESKXCPDA---ASTCTLTWCTGSGGLVLCQTKHFFWADGTSC 326

QY 61 GVEKWCCKGRC 71
DB 327 GEGKWCINGKC 337

Search completed: March 13, 2004, 07:43:16
Job time : 8.02604 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OK protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 10.3911 Seconds
(without alignments)
1930.574 Million cell updates/sec

Title: US-09-836-712-2_COPY_324_394

Perfect score: 399
Sequence: 1 PGLYISANQCRVAFGPKAV,.....VPLLDGTECGVEKWSKGRG 71

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : A_Geneseq_29Jan04:*
- 1: Geneseq1980s:*
 - 2: Geneseq1950s:*
 - 3: Geneseq2000s:*
 - 4: Geneseq2001s:*
 - 5: Geneseq2002s:*
 - 6: Geneseq2003as:*
 - 7: Geneseq2003bs:*
 - 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	399	100.0	933	5 AAU79217	Human ADA
2	399	100.0	1297	6 AAO16618	Human von
3	399	100.0	1353	5 AAe24449	Human von
4	399	100.0	1353	6 AAO16617	Human von
5	399	100.0	1416	5 ABB04153	Human ADA
6	399	100.0	1427	5 AAe24450	Human von
7	399	100.0	1427	7 ADD94038	Human agg
8	399	100.0	1445	5 ABB98125	Human PMW
9	348.5	87.3	984	6 ABR96171	Human NOV
10	256	64.2	1120	4 AAG63829	Amino aci
11	256	64.2	1120	5 ABG74113	Human mat
12	242	60.7	364	5 AAU79642	Human Hsa
13	242	60.7	364	5 ABP43989	Procollag
14	242	60.7	365	7 ADD94035	Human Hsa
15	242	60.7	1323	6 AAO16620	Human von
16	242	60.7	1378	6 AAO16619	Human von
17	181	45.4	351	7 ADD44596	Polypepi
18	181	45.4	566	4 AAe10349	Human ADA
19	181	45.4	745	4 AAe10347	Human ADA
20	181	45.4	756	6 ABP71968	Human agg
21	181	45.4	779	6 ABP71969	Human agg
22	181	45.4	807	6 ABP71967	Human agg
23	181	45.4	891	4 AAe10350	Human ADA
24	181	45.4	1057	6 ABP71970	Human agg
25	181	45.4	1095	6 ABR40094	Human ADA

26	181	45.4	1122	6 ABP71971	Human agg
27	181	45.4	1145	6 ABP71974	Human agg
28	181	45.4	1428	5 AAU72894	Human met
29	152	38.1	738	5 AAU97643	Human Hsa
30	152	38.1	738	7 ADD94036	Human Hsa
31	148	37.1	969	5 AAe21003	Human pro
32	148	37.1	980	5 AAe21000	Human pro
33	148	37.1	1035	5 ABB98128	Human PMW
34	148	37.1	1094	5 AAU72900	Human met
35	148	37.1	1207	6 ABR40096	Human ADA
36	148	37.1	1213	5 AAe21002	Human pro
37	148	37.1	1224	5 AAe20999	Human pro
38	148	37.1	1266	7 ADE47762	Human NOV
39	144	36.1	758	6 ABR96142	Human NOV
40	144	36.1	762	6 ABR96143	Human NOV
41	144	36.1	1201	3 AAB21254	Human met
42	144	36.1	1205	4 AAB73550	Human ADA
43	144	36.1	1205	6 ABR96141	Human NOV
44	143.5	36.0	730	6 ABU08955	Human ADA
45	143.5	36.0	1593	6 ABR96136	Human NOV

ALIGNMENTS

RESULT 1
AAU79217
ID AAU79217 standard; protein; 933 AA.
XX

AC AAU79217;

DT 15-JUL-2002 (first entry)

XX Human ADAM-TS-like protein.

DE Human ADAM-TS-like protein.

XX Human; ADAM-TS-like protein; cardiovascular disorder; angina;

KW vascular system; congestive heart failure; myocardial infarction;

KW ischaemic heart disease; arrhythmia; hypertensive vascular disease;

KW secondary arterial hypertension; peripheral vascular disease; embolism;

KW chronic peripheral arterial occlusive disease; acute arterial thrombosis;

KW inflammatory vascular disorder; chronic obstructive pulmonary disease;

KW liver disorder.

XX Homo sapiens.

OS Homo sapiens.

XX WO200226999-A2.

PN 04-APR-2002.

XX 26-SEP-2001; 2001WO-EP011124.

 XX 28-SEP-2000; 2000US-0235881P. || PR 25-JUL-2001; 2001US-0307393P. | XX |
XX (FARB) BAYER AG.	PA
XX Xiao Y;	PI
XX WPI; 2002-383274/41.	DR
XX N-PSDB; ABK49356.	DR
XX New purified human ADAM-TS-like protein, useful for identifying	XX
PT modulators of protein activity for treating cardiovascular or liver	PT
disorder or chronic obstructive pulmonary disease.	PT
XX Claim 25; Fig 2; 106pp; English.	XX
PS The invention relates to a human ADAM-TS-like protein and the	PS
XX polynucleotide encoding it. The protein of the invention is useful for	XX
CC treating cardiovascular disorders including diseases of the heart and	CC
CC vascular system, such as congestive heart failure, myocardial infarction,	CC
CC ischaemic heart diseases (e.g., stable angina, unstable angina), atrial	CC
CC and ventricular arrhythmia, hypertensive vascular diseases (e.g., all	CC

CC kinds of secondary arterial hypertension), and peripheral vascular
CC diseases (e.g., chronic peripheral arterial occlusive disease, acute
CC arterial thrombosis and embolism, inflammatory vascular disorders, etc),
CC liver disorders and chronic obstructive pulmonary disease. The sequences
CC are useful in diagnostic assays for detecting diseases and abnormalities
CC or susceptibility to diseases and abnormalities related to the presence
CC of mutations in the nucleic acid sequences which encode the protein. The
CC sequences are also useful for modulating ADAM-TS-like protein activity in
CC a disease condition. This sequence represents the human ADAM-TS-like
CC protein

XX SQ Sequence 933 AA;

Query Match 100.0%; Score 399; DB 5; Length 933;
Best Local Similarity 100.0%; Pred. No. 2e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
|||
Db 304 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 363
|||

QY 61 GVEKWCCKGRC 71
|||

Db 364 GVEKWCCKGRC 374
|||

RESULT 2

AAO16618

ID AAO16618 standard; protein; 1297 AA.

XX AAO16618;

AC AAO16618;

XX 15-MAY-2003 (first entry)

DT Human von Willebrand factor (vWF)-cleaving enzyme-related protein #5.

XX Human von Willebrand factor (vWF)-cleaving enzyme-related protein #5.

DE Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

XX von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

XX platelet thrombosis; stenosis.

XX Homo sapiens.

OS Homo sapiens.

XX WO200288366-A1.

PN 07-NOV-2002.

PD 25-APR-2002; 2002WO-JP004141.

PF 25-APR-2001; 2001JP-00128342.

XX 27-JUL-2001; 2001JP-00227510.

PR 28-SEP-2001; 2001JP-00302977.

XX 25-JAN-2002; 2002JP-00017596.

PR (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX WPI; 2003-120479/11.

DR N-PSDB; ABT32584.

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and

PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for

PT developing drugs for e.g. myocardial infarction and cerebral infarction.

XX Claim 4; Page 92-101; 144pp; Japanese.

PS The invention comprises the amino acid and coding sequence of a von

XX Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of

CC the invention are useful in the diagnosis and treatment of

CC thrombocytopenic purpura, and in developing drugs for myocardial

CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,

CC and stenosis. The present amino acid sequence represents a human von

CC Willebrand factor (vWF)-cleaving enzyme-related protein

XX SQ Sequence 1297 AA;

Query Match 100.0%; Score 399; DB 6; Length 1297;

Best Local Similarity 100.0%; Pred. No. 2.9e-40;

Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
|||
Db 227 PGLYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 286
|||

QY 61 GVEKWCCKGRC 71
|||

Db 287 GVEKWCCKGRC 297
|||

RESULT 3

AAE24449

ID AAE24449 standard; protein; 1353 AA.

XX AAE24449;

AC AAE24449;

XX 04-OCT-2002 (first entry)

DT Human Von Willebrand factor-cleaving protease fragment #2.

XX Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;

XX transgenic animal; immunisation; thromboembolic disease; presclampsia;

KW thrombotic thrombocytopenic purpura; TTP; Henoch-Schonlein purpura;

XX thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;

KW transgenic; anticoagulant.

XX Homo sapiens.

OS Homo sapiens.

XX WO200242441-A2.

PN 30-MAY-2002.

PD 20-NOV-2001; 2001WO-EP013391.

XX 22-NOV-2000; 2000US-00721254.

PR 12-APR-2001; 2001US-00833328.

XX (BAXT) BAXTER AG.

XX Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;

PI Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;

PI Zimmermann K, Voelkel D;

XX WPI; 2002-479950/51.

XX Novel isolated or substantially purified Von Willebrand factor-cleaving

PT protease, useful for producing preparation for therapy of thrombosis and

PT thromboembolic disease such as thrombotic thrombocytopenic purpura.

XX Claim 1; Page 64-68; 93pp; English.

XX The invention relates to an isolated or substantially pure Von Willebrand

CC factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for

CC purifying vWF which involves providing vWF-cp as a ligand, contacting a

CC solution comprising vWF with the polypeptide ligand under conditions

CC where vWF is bound to the ligand and recovering from the ligand purified

CC vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies

CC which involves immunising an animal with vWF-cp and isolating the anti-

CC vWF cp polypeptide antibodies from the animal. vWF-cp is useful for

CC producing a preparation of prophylaxis and therapy of thrombosis and

CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),

CC Henoch-Schonlein purpura, presclampsia, neonatal thrombocytopenia or

CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing

CC plasmatic or recombinantly produced vWF. The invention is useful for

CC construction expression systems and generating transgenic animals which

CC express the polypeptide in vivo. The present sequence is human vWF-cp

FF	387..439		
FT	/note="Thromspondin type I motif"		
FT	441..553		
FT	/note="Cysteine rich region"		
FT	554..687		
FT	/note="Spacer"		
FT	688..743		
FT	/note="Thromspondin type I motif"		
FT	744..805		
FT	/note="Thromspondin type I motif"		
FT	897..952		
FT	/note="Thromspondin type I motif"		
FT	953..1013		
FT	/note="Thromspondin type I motif"		
FT	1016..1073		
FT	/note="Thromspondin type I motif"		
FT	1075..1131		
FT	/note="Thromspondin type I motif"		
XX			
PN	WO200242441-A2.		
XX			
PD	30-MAY-2002.		
PP			
PF	20-NOV-2001; 2001WO-EP013391.		
XX			
PR	22-NOV-2000; 2000US-00721254.		
PPR	12-APR-2001; 2001US-00833328.		
XX			
XX	(BAXT) BAXTER AG.		
XX			
PI	Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;		
PI	Schefflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;		
PI	Zimmermann K, Voelkel D;		
XX			
XX	WPI; 2002-479950/51.		
DR	N-PSDB; AAD39332.		
XX			
PT	Novel isolated or substantially purified Von Willebrand factor-cleaving		
PT	protease, useful for producing preparation for therapy of thrombosis and		
PT	thromboembolic disease such as thrombotic thrombocytopenic purpura.		
XX			
PS	Claim 1; Fig 5; 93pp; English.		
XX			
XX	The invention relates to an isolated or substantially pure Von Willebrand		
CC	factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for		
CC	purifying vWF which involves providing vWF-cp as a ligand, contacting a		
CC	solution comprising vWF with the polypeptide ligand under conditions		
CC	where vWF is bound to the ligand and recovering from the ligand purified		
CC	vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies		
CC	which involves immunising an animal with vWF-cp and isolating the anti-		
CC	vWF cp polypeptide antibodies from the animal. vWF-cp is useful for		
CC	producing a preparation of prophylaxis and therapy of thrombosis and		
CC	thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),		
CC	Henoch-Schonlein purpura, preeclampsia, neonatal thrombocytopenia or		
CC	haemolytic-uraemic syndrome. vWF-cp can also be used for processing		
CC	plasmatic or recombinantly produced vWF. The invention is useful for		
CC	constructing expression systems and generating transgenic animals which		
CC	express the polypeptide in vivo. The present sequence is human vWF-cp		
CC	protein. vWF-cp gene is located on chromosome 9		
XX			
XX			
SQ	Sequence 1427 AA;		
	Query Match 100.0%; Score 399; DB 5; Length 1427;		
	Best Local Similarity 100.0%; Pred. No. 3.2e-40;		
	Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 PGLYISANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDFLDQSSCSRLVPLLDGTEC 60		
Db	301 PGLYISANEQCRVAFGPKAVACTFAREHLDMCQALSCHTDFLDQSSCSRLVPLLDGTEC 360		
QY	61 GVEKWSKGRC 71		
Db	361 GVEKWSKGRC 371		

RESULT 7	
ADD94038	
ID	ADD94038 standard; protein; 1427 AA.
XX	
AC	ADD94038;
XX	
DT	29-JAN-2004 (first entry)
XX	
DE	Human aggrecanase protein amino acid sequence.
XX	
KW	aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic; cartilage degradation; osteoarthritis; inflammatory joint disease; antiarthritic; osteopathic; antiinflammatory; aggrecanase-associated disorder; osteoarthritis; inflammatory condition; human; enzyme.
XX	
OS	Homo sapiens.
XX	
PN	US2003105313-A1.
XX	
PD	05-JUN-2003.
PF	25 JAN-2002; 2002US-00057487.
XX	
PPR	16-OCT-2001; 2001US-00978979.
XX	
RA	(AMHP) AMERICAN HOME PROD CORP.
XX	
PI	Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
DR	WPI; 2003-801251/75.
DR	N-PSDB; ADD94037.
XX	
PPT	New isolated DNA molecule encoding an aggrecanase polypeptide for producing a purified human aggrecanase protein which can be used to develop inhibitors of aggrecanase.
XX	
PS	Claim 14; SEQ ID NO 8; 24pp; English.
XX	
CCC	This invention relates to a novel human aggrecanase protein and the DNA sequence which encodes it. Aggrecan is a major extracellular component of articular cartilage. It is a proteoglycan responsible for providing cartilage with its mechanical properties of compressability and elasticity. A proteolytic activity (aggrecanase) is responsible for the cleavage of aggrecan thereby having a role in cartilage degradation associated with osteoarthritis and inflammatory joint disease. Compounds which inhibit the activity of the protein of the invention may have antiarthritic, osteopathic or antiinflammatory activity. The invention may be used to produce a purified human aggrecanase protein. The protein (or fragment) may be used to develop inhibitors of aggrecanase, using three dimensional structural analysis or computer aided drug design. A peptide which binds to aggrecanase is used to inhibit the proteolytic degradation of aggrecan. The invention may be useful for the development of therapeutics for the treatment of aggrecanase-associated disorders, such as, osteoarthritis and other inflammatory conditions. The present sequence is that of the human aggrecanase protein (full length sequence) of the invention.
XX	
SQ	Sequence 1427 AA;
	Query Match 100.0%; Score 399; DB 7; Length 1427;
	Best Local Similarity 100.0%; Pred. No. 3.2e-40;
	Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 PGLYSANEQCVRVAFGPKAVACTFARHLDWCALSCHTDPLDQSSCSRLIVFLDDTEC 60
Dd	301 PGLYSANEQCVRVAFGPKAVACTFARHLDWCALSCHTDPLDQSSCSRLIVFLDDTEC 360
QY	61 GVEKWSKGRC 71
Dd	361 GVEKWSKGRC 371

CC found to have homology with B. taurus procollagen I N-proteinase
XX Sequence 1445 AA;
SQ
Query Match 100.0%; Score 399; DB 5; Length 1445;
Best Local Similarity 100.0%; Pred. No. 3.2e-40;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 304 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 363
QY 61 GVEKWCCKGRC 71
DB 364 GVEKWCCKGRC 374
RESULT 9
ID ABR96171 standard; protein; 984 AA.
XX
AC ABR96171;
DT 15-SEP-2003 (first entry)
XX
DE Human NOV17a protein SEQ ID NO:84.
XX
KW Human; NOVX; G protein-coupled receptor; cytosolic; cardiovascular;
KW immunosuppressive; anti-HIV; antiasthmatic; antiarteriosclerotic; AIDS;
KW hypotensive; gene therapy; cardiomyopathy; atherosclerosis; hypertension;
KW congenital heart defect; aortic stenosis; atrial septal defect; neoplasm;
KW atrioventricular canal defect; pulmonary stenosis; prostate cancer;
KW uterine cancer; graft versus host disease; multiple sclerosis; GPCR;
KW acquired immunodeficiency syndrome; Crohn's disease; bronchial asthma;
KW chromosome mapping; forensic identification.
XX
OS Homo sapiens.
XX
PN WO200290568-A2.
XX
PD 14-NOV-2002.
XX
XX 02-MAY-2002; 2002WO-US014341.
XX PF
XX 03-MAY-2001; 2001US-0288935P.
XX PR
XX 07-MAY-2001; 2001US-0289087P.
XX PR
XX 08-MAY-2001; 2001US-0289620P.
XX PR
XX 09-MAY-2001; 2001US-0289621P.
XX PR
XX 09-MAY-2001; 2001US-0289817P.
XX PR
XX 09-MAY-2001; 2001US-0289818P.
XX PR
XX 11-MAY-2001; 2001US-0290194P.
XX PR
XX 14-MAY-2001; 2001US-0290753P.
XX PR
XX 15-MAY-2001; 2001US-0291189P.
XX PR
XX 16-MAY-2001; 2001US-0291243P.
XX PR
XX 18-MAY-2001; 2001US-0292001P.
XX PR
XX 21-MAY-2001; 2001US-0292374P.
XX PR
XX 22-MAY-2001; 2001US-0292587P.
XX PR
XX 23-MAY-2001; 2001US-0293107P.
XX PR
XX 24-MAY-2001; 2001US-0293589P.
XX PR
XX 25-MAY-2001; 2001US-0293747P.
XX PR
XX 29-MAY-2001; 2001US-0294110P.
XX PR
XX 30-MAY-2001; 2001US-0294434P.
XX PR
XX 14-AUG-2001; 2001US-0312192P.
XX PR
XX 17-AUG-2001; 2001US-0313173P.
XX PR
XX 17-AUG-2001; 2001US-0313187P.
XX PR
XX 12-SEP-2001; 2001US-0318728P.
XX PR
XX 12-SEP-2001; 2001US-0318744P.
XX PR
XX 15-NOV-2001; 2001US-0335910P.
XX PR
XX 28-NOV-2001; 2001US-0333891P.
XX PR
XX 28-NOV-2001; 2001US-0333942P.
XX PR
XX 03-JAN-2002; 2002US-0345776P.
XX PR
XX 04-JAN-2002; 2002US-0345220P.
XX PR
XX 01-MAY-2002; 2002US-00136071.

XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Casman SJ;
PI Edinger SR, Ellerman K, Gangolli EA, Gerlach VL, Gorman L;
PI Gunther E, Herrmann JL, Ji W, Lepley DM, Lewin DA, Li L;
PI Macdougall JR, Malyankar UM, Mezes PD, Padigara M, Patturajan M;
PI Feyman JA, Rastelli L, Rieger DK, Rothenberg ME, Shenoy SG;
PI Smithson G, Spytak KA, Stone DJ, Taupier RJ, Tchernev VT;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong H, Miller CE;
XX WPI; 2003-111987/10.
DR N-PSDB; ACF16980.
XX
XX New NOVX polypeptides and polynucleotides useful for treating or
PT preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
PT heart defects, aortic stenosis, atrial septal defect, or atrioventricular
PT canal defect.
XX
PS Claim 1; Page 173; 491pp; English.
XX
XX ACF16939 to ACF17000 encode the human G protein-coupled receptor (GPCR)
CC proteins, designated NOVX proteins, given in ABR96130 to ABR96191. The
CC NOVX sequences can have cytostatic, cardiovascular, antiasthmatic,
CC immunosuppressive, anti-HIV (human immunodeficiency virus), hypotensive
CC and antiarteriosclerotic activities, and can be used in gene therapy.
CC NOVX polypeptides can be used for treating a syndrome associated with a
CC human disease such as a pathology associated with the polypeptide. NOVX
CC polypeptides, polynucleotides and antibodies can be used for treating or
CC preventing e.g. cardiomyopathy, atherosclerosis, hypertension, congenital
CC heart defects, aortic stenosis, atrial septal defect, atrioventricular
CC canal defect, pulmonary stenosis, prostate cancer, uterine cancer,
CC neoplasm, graft versus host disease, acquired immunodeficiency syndrome
CC (AIDS), Crohn's disease, multiple sclerosis, or bronchial asthma. The
CC nucleic acid sequences may be used in chromosome mapping, identifying
CC individual from minute biological samples (tissue typing), and in
CC forensic identification of a biological sample. ACF17001 to ACF17117
CC represent PCR primers and probes for the NOVX sequences, which are used
CC in an example from the present invention
XX
SQ Sequence 984 AA;
Query Match 87.3%; Score 348.5; DB 6; Length 984;
Best Local Similarity 90.1%; Pred. No. 3.8e-34;
Matches 64; Conservative 0; Mismatches 0; Indels 7; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 513 PGLYISANEQCRVAFGPKAVAC-----DWQALUSCHTDPLDQSSCSRLVPLLDGTEC 565
QY 61 GVEKWCCKGRC 71
DB 566 GVEKWCCKGRC 576
RESULT 10
AAG63829
ID AAG63829 standard; protein; 1120 AA.
XX
XX AAG63829;
XX
XX 29-OCT-2001 (first entry)
DT
XX Amino acid sequence of a human zdtint5 polypeptide.
DE
XX Human; zdtint5; anti-angiogenic; intestinal polypeptide; wound healing;
KW extracellular matrix interaction; tumour suppression; gamete maturation;
KW immunologic recognition; gastrointestinal irradiation; chemotherapy;
KW proteolysis; apoptosis; angiogenesis; infection; cell adhesion;
KW cell fusion; cell signalling; tumour; Crohn's disease; melanoma; trauma;
KW inflammatory bowel disease; food poisoning; degenerative disease;
KW inflammation; fertility; gamete maturation; epithelial disorder.
XX

OS XX Homo sapiens.

PH XX Key Location/Qualifiers

FT XX Modified-site 172 /note= "potential N-linked glycosylation site"

FT XX Modified-site 176 /note= "potential N-linked glycosylation site"

FT XX Misc-difference 474 /note= "unspecified residue encoded by TTN"

FT XX Modified-site 485 /note= "potential N-linked glycosylation site"

FT XX Modified-site 533 /note= "potential N-linked glycosylation site"

FT XX Modified-site 560 /note= "potential N-linked glycosylation site"

FT XX Modified-site 595 /note= "potential N-linked glycosylation site"

FT XX Modified-site 635 /note= "potential N-linked glycosylation site"

XX XX WO200159112-A1.

XX XX 16-AUG-2001.

XX XX 09-FEB-2001; 2001WO-US004198.

XX XX 10-FEB-2000; 2000US-00501806.

XX XX (ZYMO) ZYMOGENETICS INC.

XX XX Holloway JL, Sheppard PO;

XX XX WPI; 2001-522477/57.

XX XX N-PSDB; AAH74765.

XX XX New anti-angiogenic intestinal polypeptides, zint5 polypeptides, which are members of disintegrin proteases, for modulating extracellular matrix interaction, tumor suppression and wound healing.

XX XX Claim 2; Page 84-89; 92pp; English.

XX XX The present sequence represents a human zint5 polypeptide. The zint5 polypeptide is an anti-angiogenic intestinal polypeptide. Zint5 is used for modulating extracellular matrix interactions. Zint5 polypeptide is useful as a tool for identifying new family members of polypeptides. Zint5 polynucleotides are useful as probes or primers to clone 5' non-coding regions of zint5 gene. Zint5 polypeptides are used for tumor suppression, gamete maturation, immunologic recognition, and growth and differentiation either working in isolation or in conjunction with other molecules in colon, small intestine, fetal lung, testis and B-cells. Zint5 polypeptides are also useful for promoting wound healing in the treatment of disorders associated with recovery after gastrointestinal irradiation, chemotherapy or antibody use, as anti-infectives, and extracellular matrix repair and remodeling. The polypeptides are also useful for modulating proteolysis, apoptosis, angiogenesis, infection, cell adhesion, cell fusion and signalling. The polypeptides are also useful for treating tumor formation, Crohn's disease, inflammatory bowel disease, food poisoning, melanoma, degenerative diseases, disorders related to immunity, inflammation, fertility, gamete maturation, immunology, trauma and epithelial disorders

XX XX Sequence 1120 AA;

SQ Query Match 64.2%; Score 256; DB 4; Length 1120;

Best Local Similarity 72.0%; Pred. No. 1.3e-22;

Matches 54; Conservative 0; Mismatches 9; Indels 12; Gaps 3;

QY 2 GLYSANEQCRVAPGPKAVACTFAREHLDY---COALSCHTDPLQSSCSRLVPLDGT 59

|||||

DB 331 GLYSANEQCHVAFGPG-----CRLLHQCAPCOALSCHTDPLQSSCSRLVPLDGT 385

QY 59 ECGVEK----WCSKG 69

|||||

Db 386 ECGVEKHGRWSSWG 400

RESULT 11

ABG74113

ID ABG74113 standard; protein; 1120 AA.

XX AC ABG74113;

XX 01-MAY-2003 (first entry)

XX Human mature zint5.

DE Human, zint5; ADAM; a disintegrin and metalloprotease; detergent; host defence; tumour; extracellular matrix repair; proteolysis; apoptosis; angiogenesis; bacterial infection; Crohn's disease; inflammatory bowel disease; food poisoning; melanoma; degenerative disease; chromosome 9q34; wound healing; TSP-1; Thrombospondin domain.

XX OS Homo sapiens.

XX Key Location/Qualifiers

PH FT Misc-difference 474 /label= UNKNOWN

FT /note= "Encoded by TTN"

FT Misc-difference 1118. 1119

FT /note= "Encoded by CTGTGAAT"

XX XX US2002142439-A1.

XX XX 03-OCT-2002.

XX XX 09-FEB-2001; 2001US-00781080.

XX XX 10-FEB-2000; 2000US-0181511P.

XX XX (HOLL/) HOLLOWAY J L.

XX XX (SHEP/) SHEPPARD P O.

XX XX (YAMA/) YAMAMOTO G.

XX XX Holloway JL, Sheppard PO, Yamamoto G;

XX WPI; 2003-174136/17.

XX N-PSDB; ABX15947, ABX15948.

XX New zint5 polypeptides, useful for diagnosing, preventing or treating tumor formation, Crohn's disease, inflammatory bowel disease, food poisoning, melanoma, degenerative disease, bacterial infection, or for wound healing.

XX Claim 2; Page 32-34; 37pp; English.

XX The invention relates to an isolated polypeptide, zint5, (a disintegrin and metalloprotease (ADAM), the gene for which is located on human chromosome 9q34) comprising the metalloprotease domain, thrombospondin-1 domains (TSP-1) and mature protein. Also included are an isolated polynucleotide encoding a fusion protein comprising a first and a second polypeptide segments (where the first polypeptide segment comprises the polypeptide with the sequence of S2, and the second polypeptide comprises one or more (TSP1)-like domains, and where the first polynucleotide segment is positioned amino-terminally to the second polynucleotide segment), an expression vector comprising the following operably linked elements; (a) a transcription promoter; (b) a DNA segment encoding a polypeptide comprising the amino acid sequences listed above; and (c) a transcription terminator), a cultured cell into which the expression vector of has been introduced, modulating extracellular matrix interactions by combining the polypeptides with cells, producing an antibody to the polypeptides and an isolated immunogenic polypeptide comprising at least 30 contiguous amino acids of mature zint5. The polypeptides and polynucleotides are useful for modulating cellular interactions. The polypeptides and antibodies to zint5 are useful for in vivo diagnostic or therapeutic applications, such as in identifying or

CC treating tissues or organs that express a corresponding anti-
 CC complementary molecule. The polypeptides are useful for diagnosing,
 CC preventing or treating tumour formation, Crohn's disease, inflammatory
 CC bowel disease, food poisoning, melanoma, degenerative disease, bacterial
 CC infections, extracellular matrix repair and remodeling, or for wound
 CC healing. The polypeptides are also useful for the study of host defence,
 CC extracellular matrix repair, proteolysis, apoptosis, angiogenesis and as
 CC a detergent. The antibodies may be used for tagging cells that express
 CC zdn5, for isolating zdn5 by affinity purification, or for diagnostic
 CC assays for determining circulating levels of zdn5 polypeptides. The
 CC present sequence represents mature zdn5

XX Sequence 1120 AA;

Query Match 64.2%; Score 256; DB 6; Length 1120;
 Best Local Similarity 72.0%; Pred. No. 1.3e-22;
 Matches 54; Conservative 0; Mismatches 9; Indels 12; Gaps 3;

QY 2 GLYYSANEQCRVAFGPKAVACTPAREHLM---CQALSCHTDPLDQSCSRLLVPLDGT 58
 |||||
 Db 331 GLYYSANEQCHVAFGPPG-----CRLHLRQCAPCQALSCHTDPLDQSCSRLLVPLDGT 385

QY 59 ECGVEK----WCSKG 69
 |||||

Db 386 ECGVEKHGRWSSWG 400

RESULT 12

AAU97642
 ID AAU97642 standard; protein; 364 AA.

AC AAU97642;

DT 27-AUG-2002 (first entry)

XX Human Hsa011374 protein.

DE Aggrecanase; osteopathic; antiarthritic; antiinflammatory; human;

KW Cartilage; osteoarthritis; inflammatory disease; enzyme; Hsa011374.

XX Homo sapiens.

XX WO200233093-A2.

XX 25-APR-2002.

XX 17-OCT-2001; 2001WO-US032458.

XX 18-OCT-2000; 2000US-0241469P.

XX (GEMY) GENETICS INST INC.

XX Racie LA, Twine NC, Agostino MJ, Wolfman NM, Morris EA;

XX WPI; 2002-454602/48.

XX N-PSDB; ABK52580.

XX Novel purified aggrecanase polypeptide useful for developing inhibitors
 PT and antibodies to the aggrecanase polypeptide, which are useful for
 PT treating aggrecanase-associated condition such as osteoarthritis.

XX Disclosure; Page 37-38; 41pp; English.

XX This invention relates to the cDNA and protein sequences of a novel human
 CC aggrecanase polypeptide. The protein of the invention may be used to
 CC inhibit the proteolytic activity of aggrecanase, or to inhibit the
 CC aggrecanase-mediated cleavage of aggrecan in cartilage. The protein of
 CC the invention is useful for developing inhibitors of aggrecanase protein.
 CC The cDNA sequence encoding the aggrecanase protein of the invention is
 CC useful for designing probes for obtaining DNA sequences encoding other
 CC aggrecanase molecules. The cDNA sequence is also useful for detecting
 CC mRNA encoding aggrecanase in a given cell population, and thus for
 CC detecting or diagnosing genetic disorders involving the aggrecanase, or

CC disorders involving cellular, organ or tissue disorders in which
 CC aggrecanase is irregularly transcribed or expressed. The DNA sequences
 CC may also be useful for preparing vectors for gene therapy applications.
 CC An inhibitor of the protein is useful in treating conditions
 CC characterised by degradation of articular cartilage, by blocking the
 CC enzyme's proteolytic activity. An aggrecanase protein inhibitor and a
 CC method for inhibition of its activity are useful for treating various
 CC aggrecanase-associated conditions including osteoarthritis and other
 CC inflammatory diseases. The present sequence represents the human
 CC Hsa011374 protein of the invention

XX Sequence 364 AA;

Query Match 60.7%; Score 242; DB 5; Length 364;
 Best Local Similarity 97.7%; Pred. No. 2e-21;
 Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 29 LDMCQALSCHTDPLDQSCSRLLVPLDGTGCGVEKWCCKGRC 71
 :|||||

Db 1 LDMCQALSCHTDPLDQSCSRLLVPLDGTGCGVEKWCCKGRC 43

RESULT 13

ABP43989

ID ABP43989 standard; protein; 364 AA.

AC ABP43989;

DT 26-FEB-2003 (first entry)

XX Procollagen I N-proteinase.

XX Neuroprotective; immunomodulator; cancer; chromosome 9p34; cytostatic;

KW anti-inflammatory; Gene therapy; nutritional supplement; wound; burn;

KW ulcer; Alzheimer's disease; Huntington's disease;

KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;

KW vulnary.

XX Bos taurus.

XX WO200231111-A2.

XX 18-APR-2002.

XX 11-OCT-2001; 2001WO-US027760.

XX 12-OCT-2000; 2000US-00687527.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

XX WPI; 2002-426278/45.

XX N-PSDB; ABQ61233.

XX New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.

XX Claim 20; SEQ ID # 892; 357bp + Sequence Listing; English.

XX The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC vulnary, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records ABP43544-

CC ABP43989 represent polypeptides encoded by polynucleotides of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pat_sequences
 XX
 SQ Sequence 364 AA;

Query Match 60.7%; Score 242; DB 5; Length 364;
 Best Local Similarity 97.7%; Pred. No. 2e-21; Indels 0; Gaps 0;
 Matches 42; Conservative 1; Mismatches 0;

QY 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTCGVEKWCCKGRC 71
 :|||||
 Db 1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTCGVEKWCCKGRC 43

RESULT 14
 ADD94035
 ID ADD94035 standard; protein; 365 AA.
 XX
 AC ADD94035;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Hsa011374 protein 1 related to aggrecanase.
 XX
 KW aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
 KW cartilage degradation; osteoarthritis; inflammatory joint disease;
 KW antiarthritic; osteopathic; antiinflammatory;
 KW aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
 KW human; Hsa011374.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 365
 FT /label= OTHER
 FT /note= "OTHER= Unknown amino acid; Encoded by TGA"
 XX
 US2003105313-A1.
 XX
 PD 05-JUN-2003.
 XX
 PF 25-JAN-2002; 2002US-00057487.
 XX
 PR 16-OCT-2001; 2001US-00978979.
 XX
 PA (AMHP) AMERICAN HOME PROD CORP.
 XX
 PI Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
 XX
 WPI; 2003-801251/75.
 DR N-PSDB; ADD94034.
 XX
 PT New isolated DNA molecule encoding an aggrecanase polypeptide for
 PT producing a purified human aggrecanase protein which can be used to
 PT develop inhibitors of aggrecanase.
 XX
 PS Disclosure; SEQ ID NO 5; 24pp; English.
 XX
 CC This invention relates to a novel human aggrecanase protein and the DNA
 CC sequence which encodes it. Aggrecan is a major extracellular component of
 CC articular cartilage. It is a proteoglycan responsible for providing
 CC cartilage with its mechanical properties of compressability and
 CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
 CC cleavage of aggrecan thereby having a role in cartilage degradation
 CC associated with osteoarthritis and inflammatory joint disease. Compounds
 CC which inhibit the activity of the protein of the invention may have
 CC antiarthritic, osteopathic or antiinflammatory activity. The invention
 CC (or fragment) may be used to produce a purified human aggrecanase protein. The protein
 CC may be used to develop inhibitors of aggrecanase, using
 CC three dimensional structural analysis or computer aided drug design. A
 CC peptide which binds to aggrecanase is used to inhibit the proteolytic

CC degradation of aggrecan. The invention may be useful for the development
 CC of therapeutics for the treatment of aggrecanase-associated disorders,
 CC such as, osteoarthritis and other inflammatory conditions. The present
 CC sequence is the amino acid sequence of the Hsa011374 protein 1 which is
 CC related to the human aggrecanase protein of the invention.
 XX
 SQ Sequence 365 AA;

Query Match 60.7%; Score 242; DB 7; Length 365;
 Best Local Similarity 97.7%; Pred. No. 2e-21; Indels 0; Gaps 0;
 Matches 42; Conservative 1; Mismatches 0;

QY 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTCGVEKWCCKGRC 71
 :|||||
 Db 1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTCGVEKWCCKGRC 43

RESULT 15
 AAO16620
 ID AAO16620 standard; protein; 1323 AA.
 XX
 AC AAO16620;
 XX
 DT 15-MAY-2003 (first entry)
 XX
 DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #7.
 XX
 KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
 KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
 KW myocardial infarction; cerebral infarction; arteriosclerosis;
 KW platelet thrombosis; stenosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200288366-A1.
 XX
 PD 07-NOV-2002.
 XX
 PF 25-APR-2002; 2002WO-JP004141.
 XX
 PR 25-APR-2001; 2001JP-00128342.
 PR 27-JUL-2001; 2001JP-00227510.
 PR 28-SEP-2001; 2001JP-00302977.
 PR 25-JAN-2002; 2002JP-00017596.
 XX
 PA (KAGA) CEMO-SERO-THERAPEUTIC RES INST.
 XX
 PI Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
 XX
 WPI; 2003-120479/11.
 DR N-PSDB; ABT32586.
 XX
 PT von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
 PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
 PT developing drugs for e.g. myocardial infarction and cerebral infarction.
 XX
 PS Claim 4; Page 112-121; 144pp; Japanese.
 XX
 CC The invention comprises the amino acid and coding sequence of a von
 CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
 CC the invention are useful in the diagnosis and treatment of
 CC thrombocytopenic purpura, and in developing drugs for myocardial
 CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
 CC and stenosis. The present amino acid sequence represents a human von
 CC Willebrand factor (vWF)-cleaving enzyme-related protein
 XX
 SQ Sequence 1323 AA;

Query Match 60.7%; Score 242; DB 6; Length 1323;
 Best Local Similarity 97.7%; Pred. No. 8.4e-21;
 Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTCGVEKWCCKGRC 71

Db 280 MDMCQALSCHTDPLDOSSCSRLVPLLDGTECGVERKWSKGR 322

Search completed: March 13, 2004, 07:39:02
Job time : 11.3911 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 5.62083 Seconds
(without alignments)

(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2 COPY 324 394

Perfect score: 399
Sequence: 1 PGLYYSANEQCRVAFGPKAV.....VELLDGTECGVEKWCCKGRC 71

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA,*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
10:	/cgn2_6/ptodata/2/pubpaa/US09S_PUBCOMB.pcp.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	399	100.0	842	14	US-10-222-334-4	Sequence 4, Appli
2	399	100.0	1416	9	US-09-836-712-2	Sequence 2, Appli
3	399	100.0	1427	14	US-10-232-334-2	Sequence 2, Appli
4	399	100.0	1427	14	US-10-232-334-2	Sequence 8, Appli
5	256	64.2	1120	9	US-09-781-080B-11	Sequence 11, Appli
6	242	60.7	365	9	US-09-978-973-5	Sequence 5, Appli
7	242	60.7	365	14	US-10-057-487-5	Sequence 5, Appli
8	181	45.4	756	14	US-10-188-869-6	Sequence 6, Appli
9	181	45.4	779	14	US-10-188-869-8	Sequence 8, Appli
10	181	45.4	807	14	US-10-188-869-4	Sequence 4, Appli
11	181	45.4	1057	14	US-10-188-869-0	Sequence 10, Appli
12	181	45.4	1122	14	US-10-188-869-13	Sequence 13, Appli
13	181	45.4	1145	14	US-10-188-869-20	Sequence 20, Appli
14	152	38.1	738	9	US-09-978-973-6	Sequence 6, Appli
15	152	38.1	738	14	US-10-057-487-6	Sequence 6, Appli

16	148	37.1	969	9	US-09-969-515-10	Sequence 10, Appli
17	148	37.1	980	9	US-09-969-515-4	Sequence 4, Appli
18	148	37.1	1213	9	US-09-969-515-8	Sequence 8, Appli
19	148	37.1	1224	9	US-09-969-515-2	Sequence 2, Appli
20	148	37.1	1266	15	US-10-210-130-124	Sequence 124, Appli
21	144	36.1	1205	14	US-10-184-890-1	Sequence 1, Appli
22	144	36.1	1205	15	US-10-295-027-528	Sequence 528, Appli
23	143.5	36.0	730	14	US-10-163-547-16	Sequence 16, Appli
24	143.5	36.0	1593	11	US-09-981-151A-30	Sequence 30, Appli
25	143.5	36.0	1593	15	US-10-295-027-1317	Sequence 1317, Appli
26	136	34.1	353	15	US-10-161-493-128	Sequence 128, Appli
27	136	34.1	353	15	US-10-161-493-130	Sequence 130, Appli
28	136	34.1	353	15	US-10-161-493-134	Sequence 134, Appli
29	136	34.1	862	14	US-10-226-560-2	Sequence 2, Appli
30	136	34.1	1081	15	US-10-120-801-46	Sequence 46, Appli
31	136	34.1	1162	15	US-10-161-493-124	Sequence 124, Appli
32	136	34.1	1221	14	US-10-240-545A-2	Sequence 2, Appli
33	135	33.8	353	15	US-10-161-493-132	Sequence 132, Appli
34	135	33.8	952	11	US-09-981-151A-8	Sequence 8, Appli
35	135	33.8	969	9	US-09-321-987B-5	Sequence 5, Appli
36	135	33.8	986	11	US-09-981-151A-2	Sequence 2, Appli
37	135	33.8	1156	13	US-10-014-070-5	Sequence 5, Appli
38	135	33.8	1203	10	US-09-373-658-5	Sequence 5, Appli
39	135	33.8	1203	11	US-09-989-687-5	Sequence 5, Appli
40	135	33.8	1211	14	US-10-301-832-4	Sequence 4, Appli
41	135	33.8	1223	13	US-10-014-070-2	Sequence 2, Appli
42	135	33.8	1224	13	US-10-217-774-4	Sequence 4, Appli
43	135	33.8	1224	14	US-10-236-616-2	Sequence 2, Appli
44	134	33.6	1189	16	US-10-311-035-10	Sequence 10, Appli
45	134	33.6	1213	9	US-09-938-330-14	Sequence 14, Appli

ALIGNMENTS

RESULT 1

```

US-10-222-334-4
; Sequence 4, Application US/10222334
; Publication No. US20030073116a1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS13 Genes and Pr
; FILE REFERENCE: UM-07288
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-4

```

```

Query Match      100.0%; Score 399; DB 14; Length 842;
Best Local Similarity 100.0%; Pred. No. 4.2e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 1 PGLYYSANEQCRVAFGPKAVACTTAREHLDMCQALSCHTDPLDQSSCSRLLVPLLDGTTC 60
Db 301 PGLYYSANEQCRVAFGPKAVACTTAREHLDMCQALSCHTDPLDQSSCSRLLVPLLDGTTC 360

Qy 61 GVEKWCCKGRC 71
Db 361 GVEKWCCKGRC 371

```

RESULT 2
US-09-836-712-2
; Sequence 2, Application US/09836712

```
; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1416
; TYPE: PRT
; ORGANISM: Human
US-09-836-712-2

Query Match      100.0%; Score 399; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 60
DB 324 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 383

QY 61 GVEKWCCKGRC 71
DB 384 GVEKWCCKGRC 394

RESULT 3
US-10-222-334-2
; Sequence 2, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS3 Genes and Proteins and Variants, and Uses Thereof
; FILE REFERENCE: UM-07298
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2

Query Match      100.0%; Score 399; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 60
DB 301 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 360

QY 61 GVEKWCCKGRC 71
DB 361 GVEKWCCKGRC 371

RESULT 4
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrecanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487

; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1416
; TYPE: PRT
; ORGANISM: Human
US-09-836-712-2

Query Match      100.0%; Score 399; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 60
DB 324 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 383

QY 61 GVEKWCCKGRC 71
DB 384 GVEKWCCKGRC 394

RESULT 3
US-10-222-334-2
; Sequence 2, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS3 Genes and Proteins and Variants, and Uses Thereof
; FILE REFERENCE: UM-07298
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,834
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2

Query Match      100.0%; Score 399; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 60
DB 301 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 360

QY 61 GVEKWCCKGRC 71
DB 361 GVEKWCCKGRC 371

RESULT 4
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrecanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487

; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PRIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PC10851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-057-487-8

Query Match      100.0%; Score 399; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 7.4e-39;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 60
DB 301 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 360

QY 61 GVEKWCCKGRC 71
DB 361 GVEKWCCKGRC 371

RESULT 5
US-09-781-080B-11
; Sequence 11, Application US/09781080B
; Patent No. US20020142439A1
; GENERAL INFORMATION:
; APPLICANT: HOLLOWAY, JIM
; APPLICANT: SHEPPARD, PAUL
; APPLICANT: YAMAMOTO, GAYLE
; TITLE OF INVENTION: Anti-Angiogenic Intestinal Peptides,
; TITLE OF INVENTION: zdots
; FILE REFERENCE: 99-82
; CURRENT APPLICATION NUMBER: US/09/781,080B
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 11
; LENGTH: 1120
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1120)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-781-080B-11

Query Match      64.2%; Score 256; DB 9; Length 1120;
Best Local Similarity 72.0%; Pred. No. 7.1e-22;
Matches 54; Conservative 0; Mismatches 9; Indels 12; Gaps 3;

QY 2 GLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTGTC 58
DB 331 GLYISANEQCRVAFGPPG-----CRHLRQAPCAQALSCHTDPLDQSSCSRLLVPLLDGTGTC 385

QY 59 ECGVEK-----WCSKG 69
DB 386 ECGVEKVHGRSSWG 400

RESULT 6
US-09-978-979-5
; Sequence 5, Application US/09978979
; Patent No. US20020151702A1
; GENERAL INFORMATION:
; APPLICANT: Racine, Lisa, A.
; APPLICANT: Twine, Natalie, C.
; APPLICANT: Agostino, Michael, J.
; APPLICANT: Wolfman, Neil
; APPLICANT: Morris, Elisabeth
; TITLE OF INVENTION: Aggrecanase Molecules
```

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Home Products Corporation
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/378,979
FILING DATE: 16-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: GI 5435p
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-978-979-5

Query Match 60.7%; Score 242; DB 9; Length 365;
Best Local Similarity 97.7%; Pred. No. 9.8e-21;
Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTGCGVEKWCCKGRC 71
Db 1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTGCGVEKWCCKGRC 43

RESULT 7
US-10-057-487-5
Sequence 5, Application US/10057487
Publication No. US20030105313A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: Aggrecanase Molecules
FILE REFERENCE: 08702.0073
CURRENT APPLICATION NUMBER: US/10/057,487
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 60/241,469
PRIOR FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: Patentin version 3.1
SEQ ID NO 5
LENGTH: 365
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: unknown amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (365)...(365)
OTHER INFORMATION: unknown amino acid
US-10-057-487-5

Query Match 60.7%; Score 242; DB 14; Length 365;
Best Local Similarity 97.7%; Pred. No. 9.8e-21;
Matches 42; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 29 LDMCQALSCHTDPLDQSSCSRLVPLLDGTGCGVEKWCCKGRC 71
Db 1 MDMCQALSCHTDPLDQSSCSRLVPLLDGTGCGVEKWCCKGRC 43

RESULT 8
US-10-188-869-6
Sequence 6, Application US/10188869
Publication No. US20030148306A1
GENERAL INFORMATION:
APPLICANT: LAVALLIE, EDWARD
APPLICANT: RACIE, LISA
APPLICANT: DI BLASIO, ELIZABETH
APPLICANT: AGOSTINO, MICHAEL
TITLE OF INVENTION: AGGREGANASE MOLECULES
FILE REFERENCE: 08702.0092-00000
CURRENT APPLICATION NUMBER: US/10/188,869
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: 60/349,133
PRIOR FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 60/303,051
PRIOR FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 756
TYPE: PRT
ORGANISM: Homo sapiens
US-10-188-869-6

Query Match 45.4%; Score 181; DB 14; Length 756;
Best Local Similarity 52.1%; Pred. No. 4.3e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;

Qy 1 PGLYSANECORVAFGPKAVACTPAR--EHLDMCQALSCHTDPLDQSSCSRLVPLLDGT 58
Db 331 PGHYSANECQQLFGMNA---TFCRNMEHL-MCAGLWCLVE--GDTCKTKLDPPLDGT 384
Qy 59 ECGVEKWCCKGRC 71
Db 385 ECGADKWCRCAGEC 397

RESULT 9
US-10-188-869-8
Sequence 8, Application US/10188869
Publication No. US20030148308A1
GENERAL INFORMATION:
APPLICANT: LAVALLIE, EDWARD
APPLICANT: RACIE, LISA
APPLICANT: DI BLASIO, ELIZABETH
APPLICANT: AGOSTINO, MICHAEL
TITLE OF INVENTION: AGGREGANASE MOLECULES
FILE REFERENCE: 08702.0092-00000
CURRENT APPLICATION NUMBER: US/10/188,869
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: 60/349,133
PRIOR FILING DATE: 2002-01-16
PRIOR APPLICATION NUMBER: 60/303,051
PRIOR FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 34
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 8
LENGTH: 779
TYPE: PRT
ORGANISM: Homo sapiens
US-10-188-869-8

Query Match 45.4%; Score 181; DB 14; Length 779;

Best Local Similarity 52.1%; Pred. No. 4.4e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;
QY 1 PGLYSSANEQCRVAFGPKAVACTFAR--EHLDMCOALSCHTDPLDQSSCSRLVPLDGT 58
Db 331 PGMHYSANEQCCILFGMNA---TFCRNWEHL-MCAGLWCLVE--GDTCKTKLDPPPLDGT 384
QY 59 ECGVEKWCCKGRC 71
Db 385 ECGADKWCRCAGEC 397

RESULT 10
US-10-188-869-4
; Sequence 4, Application US/10188869
; Publication No. US20030148306A1
; GENERAL INFORMATION:
; APPLICANT: LAVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; APPLICANT: DIBLASIO, ELIZABETH
; APPLICANT: AGOSTINO, MICHAEL
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0092-00000
; CURRENT APPLICATION NUMBER: US/10/188,869
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/349,133
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/303,051
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-869-4

Query Match 45.4%; Score 181; DB 14; Length 807;
Best Local Similarity 52.1%; Pred. No. 4.6e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;
QY 1 PGLYSSANEQCRVAFGPKAVACTFAR--EHLDMCOALSCHTDPLDQSSCSRLVPLDGT 58
Db 221 PGMHYSANEQCCILFGMNA---TFCRNWEHL-MCAGLWCLVE--GDTCKTKLDPPPLDGT 274
QY 59 ECGVEKWCCKGRC 71
Db 275 ECGADKWCRCAGEC 287

RESULT 11
US-10-188-869-10
; Sequence 10, Application US/10188869
; Publication No. US20030148306A1
; GENERAL INFORMATION:
; APPLICANT: LAVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; APPLICANT: DIBLASIO, ELIZABETH
; APPLICANT: AGOSTINO, MICHAEL
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0092-00000
; CURRENT APPLICATION NUMBER: US/10/188,869
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/349,133
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/303,051
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 10
; LENGTH: 1057
; TYPE: PRT
; ORGANISM: Homo sapiens

US-10-188-869-10
Query Match 45.4%; Score 181; DB 14; Length 1057;
Best Local Similarity 52.1%; Pred. No. 6.1e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;
QY 1 PGLYSSANEQCRVAFGPKAVACTFAR--EHLDMCOALSCHTDPLDQSSCSRLVPLDGT 58
Db 401 PGMHYSANEQCCILFGMNA---TFCRNWEHL-MCAGLWCLVE--GDTCKTKLDPPPLDGT 454
QY 59 ECGVEKWCCKGRC 71
Db 455 ECGADKWCRCAGEC 467
RESULT 12
US-10-188-869-13
; Sequence 13, Application US/10188869
; Publication No. US20030148306A1
; GENERAL INFORMATION:
; APPLICANT: LAVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; APPLICANT: DIBLASIO, ELIZABETH
; APPLICANT: AGOSTINO, MICHAEL
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0092-00000
; CURRENT APPLICATION NUMBER: US/10/188,869
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/349,133
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/303,051
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1122
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-869-13

Query Match 45.4%; Score 181; DB 14; Length 1122;
Best Local Similarity 52.1%; Pred. No. 6.6e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;
QY 1 PGLYSSANEQCRVAFGPKAVACTFAR--EHLDMCOALSCHTDPLDQSSCSRLVPLDGT 58
Db 466 PGMHYSANEQCCILFGMNA---TFCRNWEHL-MCAGLWCLVE--GDTCKTKLDPPPLDGT 519
QY 59 ECGVEKWCCKGRC 71
Db 520 ECGADKWCRCAGEC 532

RESULT 13
US-10-188-869-20
; Sequence 20, Application US/10188869
; Publication No. US20030148306A1
; GENERAL INFORMATION:
; APPLICANT: LAVALLIE, EDWARD
; APPLICANT: RACIE, LISA
; APPLICANT: DIBLASIO, ELIZABETH
; APPLICANT: AGOSTINO, MICHAEL
; TITLE OF INVENTION: AGGREGANASE MOLECULES
; FILE REFERENCE: 08702.0092-00000
; CURRENT APPLICATION NUMBER: US/10/188,869
; CURRENT FILING DATE: 2002-07-05
; PRIOR APPLICATION NUMBER: 60/349,133
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/303,051
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 20

```

; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-869-20

Query Match 45.4%; Score 181; DB 14; Length 1145;
Best Local Similarity 52.1%; Pred. No. 6.7e-13;
Matches 38; Conservative 7; Mismatches 20; Indels 8; Gaps 4;

QY 1 PGLYSANEQCRVAFGPKAVACTFAR--EHLDMQALSCHTDPLDQSSCSLLVPLLDGT 58
DB 466 PGMHYSANEQCCQLFGMNA---TPCRNWEHL-MCAGLWCLVE--GDTSCRTKLDPPLDGT 519
QY 59 ECGVEKWSKGRC 71
DB 520 ECGADKWRAGEC 532

RESULT 14
US-09-978-979-6
; Sequence 6, Application US/09978979
; Patent No. US20020151702A1
; GENERAL INFORMATION:
; APPLICANT: Racie, Lisa, A.
; Twine, Natalie, C.
; Agostino, Michael, J.
; Wolfman, Neil
; Morris, Elisabeth
; TITLE OF INVENTION: Aggrecanase Molecules
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: American Home Products Corporation
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/978,979
; FILING DATE: 16-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/60/241,469
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Kapinos, Ellen J.
; REGISTRATION NUMBER: 32,345
; REFERENCE/DOCKET NUMBER: GI 5435p
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 660-5000
; TELEFAX: (973) 683-4117
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 738 amino acids
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-978-979-6

Query Match 38.1%; Score 152; DB 9; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLYSYANEQCRVAFGPKAVACTFAREHL 29
DB 153 GLYSYANEQCRVAFGPKAVACTFAREHL 180
```

Search completed: March 13, 2004, 08:07:04
Job time : 6.62083 secs

```

RESULT 15
US-10-057-487-6
; Sequence 6, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrecanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 6
; LENGTH: 738
; TYPE: PRT
; ORGANISM: homo sapien
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (43)..(43)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (192)..(192)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (255)..(255)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (258)..(258)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (374)..(374)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (397)..(397)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (452)..(452)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (475)..(475)
; OTHER INFORMATION: unknown amino acid
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (487)..(487)
; OTHER INFORMATION: unknown amino acid
US-10-057-487-6

Query Match 38.1%; Score 152; DB 14; Length 738;
Best Local Similarity 100.0%; Pred. No. 1.2e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GLYSYANEQCRVAFGPKAVACTFAREHL 29
DB 153 GLYSYANEQCRVAFGPKAVACTFAREHL 180
```

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:34:04 ; Search time 2.99531 Seconds
(without alignments)
1223.727 Million cell updates/sec

Title: US-09-836-712-2_COPY_324_394

Perfect score: 399

Sequence: 1 PGLYSANEQKRVAFGPXAV.....VPLLDGTEGVERKWCCKGR 71

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/2/iaa/5A COMB.pcp.*
- 2: /cgn2_6/prodata/2/iaa/5B COMB.pcp.*
- 3: /cgn2_6/prodata/2/iaa/6A COMB.pcp.*
- 4: /cgn2_6/prodata/2/iaa/6B COMB.pcp.*
- 5: /cgn2_6/prodata/2/iaa/PCITUS COMB.pcp.*
- 6: /cgn2_6/prodata/2/iaa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135	33.8	1211	4 US-09-491-522-5	Sequence 5, Appli
2	135	33.8	1224	4 US-09-930-872-4	Sequence 4, Appli
3	124.5	31.2	997	4 US-09-369-364A-7	Sequence 7, Appli
4	124.5	31.1	1205	4 US-09-491-522-11	Sequence 11, Appli
5	123	30.8	874	4 US-09-369-364A-15	Sequence 15, Appli
6	122	30.6	930	4 US-09-369-364A-2	Sequence 2, Appli
7	122	30.6	930	4 US-09-122-126B-15	Sequence 15, Appli
8	122	30.6	930	4 US-09-634-286A-15	Sequence 15, Appli
9	121.5	30.5	551	4 US-09-130-491-16	Sequence 16, Appli
10	121.5	30.5	727	4 US-09-445-023A-12	Sequence 12, Appli
11	120.5	30.2	727	4 US-09-445-023A-1	Sequence 1, Appli
12	119.5	29.9	608	4 US-09-130-491-13	Sequence 13, Appli
13	119.5	29.9	949	4 US-09-568-559-2	Sequence 2, Appli
14	119.5	29.9	967	4 US-09-130-491-2	Sequence 2, Appli
15	117	29.3	518	4 US-09-369-364A-22	Sequence 22, Appli
16	111	27.8	566	4 US-09-491-522-7	Sequence 7, Appli
17	111	27.8	2165	4 US-09-800-729-155	Sequence 155, App
18	109	27.3	1882	4 US-09-369-364A-13	Sequence 13, Appli
19	107.5	26.9	438	4 US-09-963-791-22	Sequence 22, Appli
20	107.5	26.9	589	4 US-09-963-791-12	Sequence 12, Appli
21	107.5	26.9	757	4 US-09-963-791-24	Sequence 24, Appli
22	107.5	26.9	837	4 US-09-122-126B-2	Sequence 2, Appli
23	107.5	26.9	837	4 US-09-634-286A-2	Sequence 2, Appli
24	107.5	26.9	908	4 US-09-963-791-2	Sequence 2, Appli
25	96.5	24.2	1081	4 US-09-369-364A-17	Sequence 17, Appli
26	87.5	21.9	481	4 US-09-130-491-8	Sequence 8, Appli
27	77.5	19.4	905	4 US-09-369-364A-9	Sequence 9, Appli

28	66.5	16.7	859	4 US-09-369-364A-5	Sequence 5, Appli
29	65.5	16.4	269	4 US-09-252-991A-31516	Sequence 31516, A
30	65.5	16.4	585	4 US-09-252-991A-32406	Sequence 32406, A
31	62	15.5	1843	3 US-09-413-814-50	Sequence 50, Appli
32	61	15.3	584	1 US-08-313-288B-17	Sequence 17, Appli
33	60.5	15.2	360	4 US-09-417-039-4	Sequence 4, Appli
34	60	15.0	116	1 US-08-139-862-4	Sequence 4, Appli
35	60	15.0	269	1 US-08-203-716-3	Sequence 3, Appli
36	60	15.0	269	1 US-08-440-179-3	Sequence 3, Appli
37	60	15.0	269	2 US-08-432-693-11	Sequence 11, Appli
38	60	15.0	269	2 US-08-599-895-11	Sequence 11, Appli
39	60	15.0	269	2 US-08-483-806-2	Sequence 2, Appli
40	60	15.0	269	3 US-09-211-290-11	Sequence 11, Appli
41	60	15.0	269	3 US-09-030-613-11	Sequence 11, Appli
42	60	15.0	269	3 US-09-322-676-11	Sequence 11, Appli
43	60	15.0	269	3 US-09-039-657-3	Sequence 3, Appli
44	60	15.0	269	3 US-08-748-547-4	Sequence 4, Appli
45	60	15.0	269	3 US-09-466-036A-11	Sequence 11, Appli

ALIGNMENTS

RESULT 1

US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-433-5556
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1211 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-5

Query Match 33.8%; Score 135; DB 4; Length 1211;
Best Local Similarity 45.1%; Pred. No. 5e-08;
Matches 32; Conservative 5; Mismatches 32; Indels 2; Gaps 2;

```
QY      1 PGLVYSANEQCVRNAGFGKAVACTFAREHLD-MCQALSCHTDPLDOSSCSRLLVPLLDTGE 59  
        ||| : |::|| ::||: |:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:  
Db     464 PGVLVDYSHQRCLRGYGAYSATF-----EDMDNVCHTLWCSVG---TTCHKSLDAAYDGTR 515  
  
QY      60 CGVEKWCKSKGRC 71
```

```

RESULT 5
US-09-369-364A-15
; Sequence 15, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids En
; FILE REFERENCES: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369

```

Query Match 30.6%; Score 122; DB 4; Length 930;
Best Local Similarity 33.8%; Pred. No. 1.4e-06;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 487 PGQTYDATQCNTLTFGPEYSVC-----PGMDVCAWLCAVVRQGMVCLTKKLPAVEGTPC 542
QY 61 GVEKWCCKGRC 71
DB 543 GKGRICLQGRK 553
RESULT 8
US-09-634-286A-15
; Sequence 15, Application US/09634286A
; Patent No. 6521436
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909A
; CURRENT APPLICATION NUMBER: US/09/634,286A
; CURRENT FILING DATE: 2000-08-09
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-634-286A-15
Query Match 30.6%; Score 122; DB 4; Length 930;
Best Local Similarity 33.8%; Pred. No. 1.4e-06;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 487 PGQTYDATQCNTLTFGPEYSVC-----PGMDVCAWLCAVVRQGMVCLTKKLPAVEGTPC 542
QY 61 GVEKWCCKGRC 71
DB 543 GKGRICLQGRK 553
RESULT 9
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-16
Query Match 30.5%; Score 121.5; DB 4; Length 551;
Best Local Similarity 36.6%; Pred. No. 9.1e-07;
Matches 26; Conservative 4; Mismatches 38; Indels 3; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60

Query Match 30.8%; Score 123; DB 4; Length 874;
Best Local Similarity 35.2%; Pred. No. 1e-06;
Matches 25; Conservative 8; Mismatches 34; Indels 4; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 350 PGLYINNVKQCELIFGSGVCPYMMQ-----CRLWNCNVDAHKGCKTQHTFPWADGTEC 405
QY 61 GVEKWCCKGRC 71
DB 406 EPCKHCKKFGFC 416
RESULT 6
US-09-369-364A-2
; Sequence 2, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 930
; TYPE: PRT
; ORGANISM: mus musculus ADAMTS-5
US-09-369-364A-2
Query Match 30.6%; Score 122; DB 4; Length 930;
Best Local Similarity 33.8%; Pred. No. 1.4e-06;
Matches 24; Conservative 11; Mismatches 32; Indels 4; Gaps 1;
QY 1 PGLYISANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTEC 60
DB 487 PGQTYDATQCNTLTFGPEYSVC-----PGMDVCAWLCAVVRQGMVCLTKKLPAVEGTPC 542
QY 61 GVEKWCCKGRC 71
DB 543 GKGRVCLQGRK 553
RESULT 7
US-09-122-126B-15
; Sequence 15, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGREGAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 15
; LENGTH: 930
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-122-126B-15

Db 62 PGTLYDANRQCQFTFGESKHCPPDA---ASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSC 118
QY 61 GVEKWCCKGRC 71
Db 119 GEGKWCVSGKC 129

RESULT 10
US-09-445-023A-12
; Sequence 12, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-445-023A-12

Query Match 30.5%; Score 121.5; DB 4; Length 727;
Best Local Similarity 36.6%; Pred. No. 1.2e-06;
Matches 26; Conservative 4; Mismatches 38; Indels 3; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTSC 60
Db 238 PGTLYDANRQCQFTFGESKHCPPDA---ASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSC 294

QY 61 GVEKWCCKGRC 71
Db 295 GEGKWCVSGKC 305

RESULT 11
US-09-445-023A-1
; Sequence 1, Application US/09445023A
; Patent No. 6565858
; GENERAL INFORMATION:
; APPLICANT: Hirose, Kunitaka
; APPLICANT: Inoguchi, Eiji
; APPLICANT: Hakozaaki, Michinori
; APPLICANT: Ishioka, Keiko
; APPLICANT: Ishida, Yukako
; APPLICANT: Matsushima, Kouji
; APPLICANT: Kuno, Kouji
; TITLE OF INVENTION: Human ADAMTS-1 protein, gene encoding the same, pharmaceutical
; TITLE OF INVENTION: composition and method of immunologically analyzing human ADAMTS
; FILE REFERENCE: Q57092
; CURRENT APPLICATION NUMBER: US/09/445,023A
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: JP 9-160422
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 727
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-445-023A-1

Query Match 30.2%; Score 120.5; DB 4; Length 727;
Best Local Similarity 36.6%; Pred. No. 1.6e-06;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTSC 60
Db 238 PGTLYDANRQCQFTFGESKHCPPDA---ASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSC 294

QY 61 GVEKWCCKGRC 71
Db 295 GEGKWCINGKC 305

RESULT 12
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6418974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match 29.9%; Score 119.5; DB 4; Length 608;
Best Local Similarity 36.6%; Pred. No. 1.8e-06;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLVPLLDGTSC 60
Db 173 PGTLYDANRQCQFTFGESKHCPPDA---ASTCTTLWCTGTSGGLLVQCQTKHFPWADGTSC 229

QY 61 GVEKWCCKGRC 71
Db 230 GEGKWCINGKC 240

RESULT 13
US-09-568-559-2
; Sequence 2, Application US/09568559
; Patent No. 6649377
; GENERAL INFORMATION:
; APPLICANT: Klonowski, Paul
; APPLICANT: Allard, John
; APPLICANT: Heller, Renu
; APPLICANT: Van Wart, Harold
; TITLE OF INVENTION: Human Aggrecanase and Nucleic Acid
; TITLE OF INVENTION: Compositions Encoding the Same
; FILE REFERENCE: ROCH-002
; CURRENT APPLICATION NUMBER: US/09/568,559
; CURRENT FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: 60/133,343
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 949
; TYPE: PRT
; ORGANISM: human
US-09-568-559-2

Query Match 29.9%; Score 119.5; DB 4; Length 949;
Best Local Similarity 36.6%; Pred. No. 2.9e-06;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTEC 60
DB 460 PGTSYDANRCQFTFGEDSKHCPDA---ASTCSTLWCTGTSGGLVLCQTKHFPWADGTSC 516

QY 61 GVEKWCCKGRC 71
DB 517 GEGKWCINGKC 527

RESULT 14
US-09-130-491-2
; Sequence 2, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 967
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-130-491-2

Query Match 29.9%; Score 119.5; DB 4; Length 967;
Best Local Similarity 36.6%; Pred. No. 3e-06;
Matches 26; Conservative 3; Mismatches 39; Indels 3; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTEC 60
DB 478 PGTSYDANRCQFTFGEDSKHCPDA---ASTCSTLWCTGTSGGLVLCQTKHFPWADGTSC 534

QY 61 GVEKWCCKGRC 71
DB 535 GEGKWCINGKC 545

RESULT 15
US-09-369-364A-22
; Sequence 22, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-5
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (99)
; OTHER INFORMATION: Xaa = M
US-09-369-364A-22

Query Match 29.3%; Score 117; DB 4; Length 518;
Best Local Similarity 33.8%; Pred. No. 3e-06;
Matches 24; Conservative 9; Mismatches 34; Indels 4; Gaps 1;

QY 1 PGLYYSANEQCRVAFGPKAVACTFAREHLDWCQALSCHTDPLDQSSCSRLLVPLLDGTEC 60
DB 75 PGQTYDATQCNCNLTFFGPDYSVC-----PGXDVCARLWCVAVRQGQMVCLTKLPAVEGTTC 130

QY 61 GVEKWCCKGRC 71
DB 131 GKGRICLQKGC 141

Search completed: March 13, 2004, 07:46:07
Job time : 3.99531 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:33:19 ; Search time 2.23333 Seconds

(without alignments)
2756.533 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473

Perfect score: 365

Sequence: 1 WSWGPRSPCSRSCGGVVT.....EMCNTQACEKTOLEFMSQQC 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: piri:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	183	50.1	1205	T18517	procollagen N-endo
2	169.5	46.4	1074	JC5928	semaphorin F precu
3	162	44.4	1172	TSHUP2	thrombospondin 2 p
4	161	44.1	1170	TSHUP1	thrombospondin 1 p
5	161	44.1	1170	A40358	thrombospondin 1 p
6	160	43.8	1178	A39804	thrombospondin pre
7	156	42.7	1172	A42587	thrombospondin 2 p
8	155.5	42.6	550	T47158	hypothetical prote
9	155.5	42.6	951	T00017	gene ADAMTS-1 prot
10	154.5	42.3	837	T00355	hypothetical prote
11	150.5	41.2	2165	T21371	hypothetical prote
12	145.5	39.9	469	S29126	properdin precursor
13	144.5	39.6	1444	T18856	angiogenesis inhib
14	140.5	38.5	1558	C89114	protein C37C3.6a [
15	140.5	38.5	2167	T34395	hypothetical prote
16	135.5	37.1	437	S05478	properdin - mouse
17	135	37.0	1572	T00027	brain-specific ang
18	129	35.3	843	A27340	complement C7 prec
19	129	35.3	951	T00260	hypothetical prote
20	126	34.5	736	T19366	hypothetical prote
21	122	33.4	1584	T00036	brain-specific ang
22	120	32.9	984	T00326	hypothetical prote
23	120	32.9	1522	T00028	'brain-specific ang
24	119.5	32.7	585	T46686	complement compone
25	117.5	32.2	957	T15976	hypothetical prote
26	113.5	31.1	690	T16761	hypothetical prote
27	113	31.0	590	T46687	complement compone
28	112.5	30.8	919	T32541	unc-5 protein - Ca
29	112.5	30.8	947	B44294	unc-5 protein, lon

30 111.5 30.5 788 2 T25061
31 109 29.9 584 1 C8HUA
32 106 29.0 1059 2 T22545
33 103 28.2 654 2 T29247
34 101 27.7 651 2 T13477
35 99.5 27.3 591 1 C8HUB
36 98 26.8 206 2 A45517
37 98 26.6 712 2 A45638
38 97 26.6 860 2 T16892
39 95.5 26.2 934 1 A34372
40 92.5 25.3 1360 2 T33922
41 91 24.9 2761 2 T21064
42 90.5 24.8 803 2 A47723
43 90 24.7 254 2 T15952
44 88.5 24.2 724 2 A48569
45 84.5 23.2 485 2 A60610

ALIGNMENTS

RESULT 1

T18517

procollagen N-endopeptidase (EC 3.4.24.14) I - bovine

N/Alternate names: procollagen N-proteinase

C/Species: Bos primigenius taurus (cattle)

C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C/Accession: T18517

R/Colige, A.; Nusgens, B.V.; Lapiere, C.M.

submitted to the EMBL Data Library, February 1996

A/Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.

A/Reference number: Z18941

A/Accession: T18517

A/Status: preliminary; translated from GB/EMBL/DBDJ

A/Molecule type: mRNA

A/Residues: 1-1205 <COL>

A/Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CAA65253.1

A/Experimental source: skin

C/Genetics:

A/Gene: PC I-NP

C/Function:

A/Description: catalyzes cleavage of the propeptides of type I and II collagens prior t

C/Keywords: hydrolase; metalloproteinase

Query Match 50.1%; Score 183; DB 2; Length 1205;
Best Local Similarity 46.9%; Pred. No. 7.9e-12;
Matches 30; Conservative 8; Mismatches 26; Indels 0; Gaps 0;
QY 1 WSWGPRSPCSRSCGGVVTTRQCNPRPAGGRACVGADLQAEMLQACEKTOLEPM 60
DB 558 WGAWSPFGSCSRCTCGTGKTRTCQDNPHDPANGRTGCGLAYDFQLCNSQDCPDALADPR 617

QY 61 SQQC 64

DB 618 EEQC 621

RESULT 2

JC5928

semaphorin F precursor - human

C/Species: Homo sapiens (man)

C/Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000

C/Accession: JC5928

R/Simmons, A.D.; Fueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.

Biochem. Biophys. Res. Commun. 242, 685-691, 1998

A/Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candi

A/Reference number: JC5928; MUID:98125554; PMID:9464278

A/Accession: JC5928

A/Status: nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 1-1074 <SIM>

A/Cross-references: GB:U52840; NID:g2772583; PIDN:AAC09473.1; PID:g2772584

A/Experimental source: brain

C;Comment: This protein disrupts normal brain development and leads to some of the featur
C;Genetics:
A;Gene: sema4
C;Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F;1-20/Domain: signal sequence #status predicted <SIG>
F;50-533/Domain: semaphorin #status predicted <SEM>
F;840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F;971-993/Domain: transmembrane #status predicted <TM>

Query Match 46.4%; Score 169.5; DB 2; Length 1074;
Best Local Similarity 37.7%; Pred. No. 2e-10;
Matches 29; Conservative 12; Mismatches 23; Indels 13; Gaps 1;

QY 1 WSWPSPSCSRSCGGVTRRRQNNPRPAGGACVADLAQEMCNQAC 50
DB 844 WSWPSPSCSRSCGGVTRRRQNNPRPAGGACVADLAQEMCNQAC 50

QY 51 ---ACEKTLQLEFMSQOC 64
DB 904 DWSECEASGVQVRAQC 920

RESULT 3
TSHUP2
C;Species: Homo sapiens (man)
C;Date: 19-May-1995 #sequence_revision 03-Aug-1995 #text_change 13-Aug-1999
C;Accession: A47379; A42173
R;LaBell, T.L.; Byers, P.H.
Genomics 17, 225-229, 1993
A;Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote
A;Reference number: A47379; MUID:94010892; PMID:8406456
A;Accession: A47379
A;Molecule type: mRNA
A;Residues: 1-1172 <LAB>
A;Cross-references: GB:L12350; NID:G307505; PIDN:AAA03703.1; PID:G307506
R;LaBell, T.L.; Milewicz, D.J.; Distche, C.M.; Byers, P.H.
Genomics 12, 421-429, 1992
A;Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression o
A;Reference number: A42173; MUID:92217961; PMID:1559694
A;Accession: A42173
A;Molecule type: mRNA
A;Residues: 560-1172 <LA2>
A;Cross-references: GB:M81339
A;Experimental source: fibroblast
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBIIP:95096)
C;Genetics:
A;Gene: GDB:THBS2; TSP2
A;Cross-references: GDB:128789; OMIM:188061
A;Map position: 6q27-6q27
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-1172/Product: thrombospondin 2 #status predicted <MAT>
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: EGF homology <EGF1>
F;652-691/Domain: EGF homology <EGF>
F;928-930/Region: cell attachment (R-G-D) motif
F;151,316,330,457,584,710,1059/Binding site: carbohydrate (Asn) (covalent) #status predi
F;167-226/Disulfide bonds: #status predicted
F;266,270/Disulfide bonds: interchain #status predicted
F;612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match 44.4%; Score 162; DB 1; Length 1172;
Best Local Similarity 48.1%; Pred. No. 1.4e-09;
Matches 25; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 WSWPSPSCSRSCGGVTRRRQNNPRPAGGACVADLAQEMCNQAC 52
DB 497 WSPSPSWACTVTCAGGIRERTVCNSPEPQYGKACVGVQERQCMCNKRC 548

RESULT 4
TSHUP1
thrombospondin 1 precursor - human
C;Species: Homo sapiens (man)
C;Date: 23-Aug-1987 #sequence_revision 03-Aug-1995 #text_change 17-Nov-2000
C;Accession: A26155; A34274; A30140; A25812; A05172; A42927
R;Lawler, J.; Hynes, R.O.
J. Cell Biol. 103, 1635-1648, 1986
A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple
A;Reference number: A26155; MUID:87057617; PMID:2430973
A;Accession: A26155
A;Molecule type: mRNA
A;Residues: 1-1170 <LAW>
A;Cross-references: GB:X04665; NID:G37137; PIDN:CAA28370.1; PID:G37138
A;Note: parts of this sequence, including the amino end of the mature protein, were det
R;LaHerty, C.D.; Gierman, T.M.; Dixit, V.M.
J. Biol. Chem. 264, 11222-11227, 1989
A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA
A;Reference number: A34274; MUID:89291870; PMID:2544587
A;Accession: A34274
A;Molecule type: DNA
A;Residues: 1-166 <LAH>
A;Cross-references: GB:J04835
R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,
J. Cell Biol. 108, 729-736, 1989
A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in t
A;Reference number: A30140; MUID:89139590; PMID:2918029
A;Accession: A30140
A;Molecule type: mRNA
A;Residues: 1-83; 'A', 85-522, 'A', 524-1170 <HEN>
A;Cross-references: EMBL:X14787; NID:G37464; PIDN:CAA32889.1; PID:G37465
A;Note: parts of this sequence, including the amino end of the mature protein, were det
R;Kobayashi, S.; Eden-McCuchan, F.; Franson, P.; Bornstein, P.
Biochemistry 25, 8418-8425, 1986
A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis
A;Reference number: A25812; MUID:87157592; PMID:3030396
A;Accession: A25812
A;Molecule type: mRNA
A;Residues: 1-83; 'A', 85-397 <KOB>
A;Cross-references: GB:M25631; NID:G538353; PIDN:AAA36741.1; PID:G538354
R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986
A;Reference number: A05172; MUID:86287276; PMID:3461443
A;Accession: A05172
A;Molecule type: mRNA
A;Residues: 1-83; 'A', 85-374, 'RC' <DIX>
A;Cross-references: GB:M14326; NID:G340005; PIDN:AAA61237.1; PID:G553801
A;Note: parts of this sequence, including the amino end of the mature protein, were det
R;Sun, X.; Skorstengaard, K.; Mosher, D.F.
J. Cell Biol. 118, 693-701, 1992
A;Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.
A;Reference number: A42927; MUID:92348511; PMID:1379247
A;Accession: A42927
A;Molecule type: protein
A;Residues: 987-1003 <SUN>
A;Note: Cys-992 is shown to have a free sulfhydryl
C;Genetics:
A;Gene: GDB:THBS1; TSP1; TSP
A;Cross-references: GDB:120438; OMIM:188060
A;Map position: 15q15-15q15
A;Introns: 23/1
A;Note: the list of introns may be incomplete
C;Complex: homotrimer, disulfide linked
C;Function:
A;Description: participates in cell migration and adhesion, and in platelet aggregation
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
C;Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim
F;1-18/Domain: signal sequence #status predicted <SIG>

F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <VWC>
F;378-429/Domain: thrombospondin type 1 repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>
F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF1>
F;650-689/Domain: EGF homology <EGF2>
F;926-928/Region: cell attachment (R-G-D) motif
F;171-232/Disulfide bonds: #status predicted
F;248-360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;270,274/Disulfide bonds: interchain #status predicted
F;610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted
F;1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 44.1%; Score 161; DB 1; Length 1170;
Best Local Similarity 51.9%; Pred. No. 1.7e-09;
Matches 27; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVVTTRRRQCNPPAFGGRACVADLQAEKMCNTQAC 52
DB 495 WGPSPWDICSVTCGGGVQKRSRLCNPTPQFGKDCVGVDTENQCNKQDC 546

RESULT 5
A40358
thrombospondin 1 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40558; B42587; S68787
R:Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.
Genomics 11, 587-600, 1991
A:Title: Characterization of the murine thrombospondin gene.
A:Reference number: A40558; MUID:92128941; PMID:1774063
A:Accession: A40558
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1170 <LAW>
A:Cross-references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; GB:M62455; GB:M62466; GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:9511867; PIDN:AAA5061
R:Bornstein, P.; Alfi, D.; Devarayalu, S.; Framson, P.; Li, P.
J. Biol. Chem. 265, 16691-16698, 1990
A:Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A:Reference number: A37905; MUID:90375546; PMID:2398070
A:Accession: A37905
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <BOR>
A:Cross-references: GB:J05605; GB:J05606; NID:G201991; PIDN:AAA40431.1; PID:G554390
R:Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.
Genomics 11, 587-600, 1991
A:Title: Characterization of the mouse thrombospondin gene and evaluation of the role of
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: A42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1152, 'P', 1154-1170 <LAH>
A:Cross-references: GB:M87276
A:Note: sequence extracted from NCBI backbone (NCBIP:81501)
R:Chen, H.; Aeschlinann, D.; Nowlen, J.; Mosher, D.F.
FEBS Lett. 387, 36-41, 1996
A:Title: Expression and initial characterization of recombinant mouse thrombospondin 1
A:Reference number: S68787; MUID:96234006; PMID:8654563
A:Accession: S68787
A:Molecule type: protein
A:Residues: 19-26, 'X', 28-37 <CHE>
C:Complex: homotrimer, disulfide linked
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von
C:Keywords: calcium binding; glycoprotein; homotrimer
F;19-1170/Product: thrombospondin 1 #status predicted <MAT>
F;317-375/Domain: von Willebrand factor type C repeat homology <THR1>
F;434-490/Domain: thrombospondin type 1 repeat homology <THR2>

F;491-547/Domain: thrombospondin type 1 repeat homology <THR3>
F;551-586/Domain: EGF homology <EGF>
F;248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 44.1%; Score 161; DB 2; Length 1170;
Best Local Similarity 51.9%; Pred. No. 1.7e-09;
Matches 27; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVVTTRRRQCNPPAFGGRACVADLQAEKMCNTQAC 52
DB 495 WGPSPWDICSVTCGGGVQKRSRLCNPTPQFGKDCVGVDTENQCNKQDC 546

RESULT 6
A39804
thrombospondin precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A39804
R:Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A:Title: Cloning and sequencing of chicken thrombospondin.
A:Reference number: A39804; MUID:91217026; PMID:2022631
A:Accession: A39804
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <LAW>
A:Cross-references: GB:M60853; NID:G212763; PIDN:AAA51437.1; PID:G212764
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology <VWC>
F;325-383/Domain: von Willebrand factor type C repeat homology <THR1>
F;386-437/Domain: thrombospondin type 1 repeat homology <THR2>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR3>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

Query Match 43.8%; Score 160; DB 1; Length 1178;
Best Local Similarity 48.1%; Pred. No. 2.2e-09;
Matches 25; Conservative 8; Mismatches 19; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVVTTRRRQCNPPAFGGRACVADLQAEKMCNTQAC 52
DB 503 WGPSPWSACTVTCGGGIRSRSLCNPSPEYGGKPCVGDTCQKQDMCNKQDC 554

RESULT 7
A42587
thrombospondin 2 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C:Accession: A42587; A39851
R:Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.
Genomics 11, 587-600, 1991
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell
A:Reference number: A42587; MUID:92147683; PMID:1371115
A:Accession: A42587
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-1172 <LAH>
A:Cross-references: GB:I07803; GB:M87275; NID:G340421; PIDN:AAA53064.1; PID:G567241
A:Note: sequence extracted from NCBI backbone (NCBIP:81502)
R:Bornstein, P.; O'Rourke, K.; Wilstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A:Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A:Reference number: A39851; MUID:91302287; PMID:1712771
A:Accession: A39851
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-873 <BOR>
A:Cross-references: GB:M64866; NID:G201994; PIDN:AAA40432.1; PID:G201995
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology <VWC>
C:Keywords: calcium binding; glycoprotein
F;319-377/Domain: von Willebrand factor type C repeat homology <THR1>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR2>

F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F:553-588/Domain: EGF homology <EGF1>
F:652-691/Domain: EGF homology <EGF>
Query Match 42.7%; Score 156; DB 2; Length 1172;
Best Local Similarity 46.2%; Pred. No. 5.9e-09;
Matches 24; Conservative 10; Mismatches 18; Indels 0; Gaps 0;
QY 1 WSWGPRSPCSRSCGGVVTTRRQCNPPRPAFGGRACVGDADLAEMCNQTQAC 52
DB 497 WSPWSPWACTVTCAGGIRSRVCSNPSPQYGGKDCVGDVTEHQMCNKRS 548
RESULT 8
T47158
hypothetical protein DKFP762C1110.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47158
R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wiemann, S.
Submitted to the Protein Sequence Database, March 2000
A:Reference number: 224379
A:Accession: T47158
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-550 <AAA>
A:Cross-references: EMBL:AL162080
A:Experimental source: adult melanoma (MeWo cell line); clone DKFP762C1110
C:Genetics:
A:Note: DKFP762C1110.1
Query Match 42.6%; Score 155.5; DB 2; Length 550;
Best Local Similarity 41.5%; Pred. No. 3.7e-09;
Matches 27; Conservative 9; Mismatches 28; Indels 1; Gaps 1;
QY 1 WSWGPRSPCSRSCGGVVTTRRQCNPPRPAFGGRACVGDADLAEMCNQTQAC-EKTQLEF 59
DB 145 WGMWPGWDCSRTCGGVQVYTWRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTF 204
QY 60 MSQOC 64
DB 205 REEQC 209
RESULT 9
T00017
gene ADAMTS-1 protein - mouse
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00017
R:Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
Genomics 46, 466-471, 1997
A:Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A:Reference number: 214055; MUID:98110583; PMID:9441751
A:Accession: T00017
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-951 <KUN>
A:Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A:Experimental source: strain 129SVJ
C:Genetics:
A:Gene: ADAMTS-1
A:Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
C:Superfamily: thrombospondin type 1 repeat homology
F:542-598/Domain: thrombospondin type 1 repeat homology <THR3>
Query Match 42.6%; Score 155.5; DB 2; Length 951;
Best Local Similarity 41.5%; Pred. No. 5.7e-09;
Matches 27; Conservative 9; Mismatches 28; Indels 1; Gaps 1;
QY 1 WSWGPRSPCSRSCGGVVTTRRQCNPPRPAFGGRACVGDADLAEMCNQTQAC-EKTQLEF 59

DB 546 WGPWPGWDCSRTCGGVQVYTWRECDNPVPKNGGKYCEGKRVYRSCNIEDCPDNNGKTF 605
QY 60 MSQOC 64
DB 606 REEQC 610
RESULT 10
T00355
hypothetical protein KIAA0688 - human
C:Species: Homo sapiens (man)
C:Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C:Accession: T00355
R:Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, I.
DNA Res. 5, 169-176, 1998
A:Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A:Reference number: Z14142; MUID:98403890; PMID:9734811
A:Accession: T00355
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-837 <ISH>
A:Cross-references: EMBL:AB014588; NID:g3327189; PIDN:BAA31663.1; PID:g3327190
A:Experimental source: brain
C:Genetics:
A:Gene: KIAA0688
C:Superfamily: thrombospondin type 1 repeat homology
F:519-575/Domain: thrombospondin type 1 repeat homology <THR3>
Query Match 42.3%; Score 154.5; DB 2; Length 837;
Best Local Similarity 43.1%; Pred. No. 6.5e-09;
Matches 28; Conservative 6; Mismatches 30; Indels 1; Gaps 1;
QY 1 WSWGPRSPCSRSCGGVVTTRRQCNPPRPAFGGRACVGDADLAEMCNQTQACEK-TQLEF 59
DB 523 WGPWPGWDCSRTCGGVQVYTWRECDNPVPKNGGKYCEGKRVYRSCNLEDCPDNNGKTF 582
QY 60 MSQOC 64
DB 583 REEQC 587
RESULT 11
T21371
hypothetical protein F25H8.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
C:Accession: T21371; T24896
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19413
A:Accession: T21371
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL>
A:Cross-references: EMBL:Z69360; PIDN:CAA93287.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone F25H8
R:Gajadaty, S.
submitted to the EMBL Data Library, February 1996
A:Reference number: Z19949
A:Accession: T24896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2165 <WIL2>
A:Cross-references: EMBL:Z69361; PIDN:CAA93288.1; GSPDB:GN00022; CESP:F25H8.3
A:Experimental source: clone T13H10
C:Genetics:
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1; 52/1; 135/2; 193/3; 216/1; 266/1; 495/2; 547/3; 584/3; 634/2; 744/1; 1;
Query Match 41.2%; Score 150.5; DB 2; Length 2165;
Best Local Similarity 43.8%; Pred. No. 3.7e-08;
Matches 28; Conservative 7; Mismatches 26; Indels 3; Gaps 1;

F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-469/Product: properdin #status experimental <NAT>
F:76-128/Domain: thrombospondin type 1 repeat homology <THR1>
F:135-191/Domain: thrombospondin type 1 repeat homology <THR2>
F:192-255/Domain: thrombospondin type 1 repeat homology <THR3>
F:256-313/Domain: thrombospondin type 1 repeat homology <THR4>
F:314-377/Domain: thrombospondin type 1 repeat homology <THR5>
F:378-440/Domain: thrombospondin type 1 repeat homology <THR6>
F:83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-mannosyl-try
F:428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.9%; Score 145.5; DB 1; Length 469;
Best Local Similarity 52.0%; Pred. No. 3.8e-08;
Matches 26; Conservative 5; Mismatches 18; Indels 1; Gaps 1;

QY 1 WSWGPRSPCSRCGGVTVTRRQCNNPRPFAFGACVADLAQEMCNTOACEKTLQLEFM 60
Db 609 WRSWG--ECSRTGGGVQKGLRDCDSFKPRNGGKYCVGRERYRSCNTQCPWDTPYR 665

QY 61 SQQC 64
Db 666 EVQC 669

RESULT 12
S29126
properdin precursor [validated] - human
N:Alternate names: factor P
C:Species: Homo sapiens (man)
C:Date: 17-Nov-2000 #sequence revision 17-Nov-2000 #text_change 17-Nov-2000
C:Accession: S29126; S16150; A05319; T45113
R:Nolan, K.F.; Kaluz, S.; Higgins, J.M.G.; Goundis, D.; Reid, K.B.M.
Biochem. J. 287, 291-297, 1992
A:Title: Characterization of the human properdin gene.
A:Reference number: S29126; MUID:93038568; PMID:1417780
A:Accession: S29126
A:Molecule type: DNA
A:Residues: 1-469 <NOLI>
A:Cross-references: EMBL:X70872; NID:G35679; PIDN:CAA50220.1; PID:G35680
R:Nolan, K.F.; Schwaebler, W.; Kaluz, S.; Dierich, M.P.; Reid, K.B.M.
Eur. J. Immunol. 21, 771-776, 1991
A:Title: Molecular cloning of the cDNA coding for properdin, a positive regulator of the
A:Reference number: S16150; MUID:91184288; PMID:2009915
A:Accession: S16150
A:Molecule type: mRNA
A:Residues: 1-456, R', 458-469 <NOL2>
A:Cross-references: EMBL:X57748
R:Reid, K.B.M.; Gagnon, J.
Mol. Immunol. 18, 949-959, 1981
A:Reference number: A05319; MUID:82195224; PMID:7341961
A:Accession: A05319
A:Molecule type: protein
A:Residues: 28-53, Q', 55-59, G', 61, I', 63, 137-138, P', 140-141, P', 143-144, X', 146-148, Y'
-238, GH', 241-245, 248-251, X', 253-257, P', 259, G', 261, XP', 265-266, X', 268-269, 280-285,
K', 407-421-427, R', 429-443, TKV', 447-448, XX', 451, RX', 454-455 <REI>
R:Westberg, J.; Nordin-Fredrikson, G.; Truedsson, L.; Sjöholm, A.G.; Uhlen, M.
Submitted to the EMBL Data Library, May 1997
A:Reference number: Z22914
A:Accession: T45112
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-54, X', 56-73, X', 75-99, W', 101-469 <WES1>
A:Cross-references: EMBL:AF005665; PIDN:AAB63280.1
A:Experimental source: genomic DNA from individual with properdin deficiency type II
A:Accession: T45113
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-60, X', 62-413, D', 415-452, XX', 455-469 <WE2>
A:Cross-references: EMBL:AF005666; PIDN:AAC51626.1
A:Experimental source: genomic DNA from individual with properdin deficiency type III
R:Hartmann, S.; Hofsteenge, J.
J. Biol. Chem. 275, 28569-28574, 2000
A:Title: Properdin, the positive regulator of complement, is highly C-mannosylated.
A:Reference number: A59360; MUID:20435812; PMID:10878002
A:Contents: annotation
A:Note: identification and location of C-mannosylation sites by mass-spectroscopy
C:Genetics:
A:Gene: GDB:PFC
A:Cross-references: GDB:120275; OMIM:312060
A:Map position: Xp11.3-Xp11.23
A:Introns: 26/1; 76/2; 135/1; 192/1; 256/1; 314/1; 378/1; 415/2
A:Complex: a mixture of homodimers, homotrimers and homotetramers
C:Function:
A:Description: protects C3 convertase (C3bBb) from rapid inactivation
A:Pathway: complement alternate pathway
C:Superfamily: human properdin precursor; thrombospondin type 1 repeat homology
C:Keywords: complement alternate pathway; glycoprotein; homodimer; homotetramer; homotri

F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-469/Product: properdin #status experimental <NAT>
F:76-128/Domain: thrombospondin type 1 repeat homology <THR1>
F:135-191/Domain: thrombospondin type 1 repeat homology <THR2>
F:192-255/Domain: thrombospondin type 1 repeat homology <THR3>
F:256-313/Domain: thrombospondin type 1 repeat homology <THR4>
F:314-377/Domain: thrombospondin type 1 repeat homology <THR5>
F:378-440/Domain: thrombospondin type 1 repeat homology <THR6>
F:83,86,139,142,145,196,199,260,263,321,324,382,385,388/Modified site: 2'-mannosyl-try
F:428/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 39.9%; Score 145.5; DB 1; Length 469;
Best Local Similarity 52.0%; Pred. No. 3.8e-08;
Matches 26; Conservative 5; Mismatches 18; Indels 1; Gaps 1;

QY 1 WSWGPRSPCSRCGGVTVTRRQCNNPRPFAFGACVADLAQEMCNTO 50
Db 139 WSWGPEPCSVTCSTGTRRRACNHPAPKCGH-CFGQAQSEACDTQ 187

RESULT 13
T18856
angiogenesis inhibitor homolog - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T18856; T24653
R:McMurray, A.
Submitted to the EMBL Data Library, July 1995
A:Reference number: Z19031
A:Accession: T18856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL>
A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN000028; CESP:CO2B4.1
A:Experimental source: clone CO2B4
R:McMurray, A.
Submitted to the EMBL Data Library, July 1995
A:Reference number: Z19917
A:Accession: T24653
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1444 <WIL2>
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN000028; CESP:CO2B4.1
A:Experimental source: clone T0705
C:Genetics:
A:Gene: CESP:CO2B4.1
A:Map position: X
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 56

Query Match 39.6%; Score 144.5; DB 2; Length 1444;
Best Local Similarity 47.2%; Pred. No. 1.2e-07;
Matches 25; Conservative 9; Mismatches 18; Indels 1; Gaps 1;

QY 1 WSWGPRSPCSRCGGVTVTRRQCNNPRPFAFGACVADLAQEMCNTOACE 53
Db 1310 WSTWSDWTPCSASCGFGVQTRDRSCSPPEPK-GQSCSGLAHQTSLCDLPACD 1361

RESULT 14
C89114
protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biol
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/projects/C_e
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA

A;Residues: 1-1558 <STO>
A;Cross-references: GB:chr_V; PIDN:AAC25867.1; PID:g3294501; GSPDB:GN00023; CESP:C37C3.6
C;Genetics:
A;Gene: C37C3.6a
A;Map position: 5

Query Match 38.5%; Score 140.5; DB 2; Length 1558;
Best Local Similarity 40.6%; Pred. No. 3.4e-07;
Matches 26; Conservative 9; Mismatches 20; Indels 9; Gaps 2;

Qy 1 WSSWGPRSPCSRCGGGVVTRRQCNNPRPAFGGRACVGDLOAEMCNTQACEKTQLEFM 60
Db 79 WGPWVPENECRSRCGGGVQLEKRCQSGD-----CTGASVRYISCNLNACE-SGTDPR 129

Qy 61 SQQC 64

Db 130 AEQC 133

RESULT 15

T34395

hypothetical protein C37C3.6b - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000

C;Accession: T34395; T34394

R;Geisel, C.; Bradshaw, H.

submitted to the EMBL Data Library, July 1996

A;Description: The sequence of C. elegans cosmid C37C3.

A;Reference number: Z21518

A;Accession: T34395

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-2167 <GE1>

A;Cross-references: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b

A;Experimental source: strain Bristol N2; clone C37C3

A;Accession: T34394

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1555, 'SKP' <GE2>

A;Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a

A;Experimental source: strain Bristol N2; clone C37C3

C;Genetics:

A;Gene: CESP:C37C3.6b; CESP:C37C3.6a

A;Map position: 5

A;Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 1556/

Query Match

Best Local Similarity 38.5%; Score 140.5; DB 2; Length 2167;

Matches 26; Conservative 9; Mismatches 20; Indels 9; Gaps 2;

Qy 1 WSSWGPRSPCSRCGGGVVTRRQCNNPRPAFGGRACVGDLOAEMCNTQACEKTQLEFM 60

Db 79 WGPWVPENECRSRCGGGVQLEKRCQSGD-----CTGASVRYISCNLNACE-SGTDPR 129

Qy 61 SQQC 64

Db 130 AEQC 133

Search completed: March 13, 2004, 07:44:39

Job time : 3.23333 secs

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 1.33333 Seconds
(without alignments)
2499.367 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473

Perfect score: 365

Sequence: 1 WSSWGRSPCRSRSCGGVVT.....EMCNCQACEXTQLEFMSQQC 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	206	56.4	1223	1 AT14_HUMAN	Q8wx58 homo sapien
2	192	52.6	1224	1 AT16_HUMAN	Q8te57 homo sapien
3	191	52.3	1935	1 AT59_HUMAN	Q9p2n4 homo sapien
4	189	51.8	1081	1 AT18_HUMAN	Q8te60 homo sapien
5	188	51.5	1906	1 AT20_MOUSE	P59511 mus musculus
6	185	50.7	1205	1 AT53_HUMAN	O15072 homo sapien
7	183	50.1	1205	1 AT32_BOVIN	P79331 b adams-2
8	181	49.6	1213	1 AT52_MOUSE	Q8c9w3 m adams-2
9	181	49.6	1911	1 AT20_HUMAN	P59510 homo sapien
10	177	48.5	950	1 AT15_HUMAN	Q8te58 homo sapien
11	177	48.5	1211	1 AT52_HUMAN	O95450 h adams-2
12	175	47.9	905	1 AT58_MOUSE	P57110 mus musculus
13	172	47.1	890	1 AT58_HUMAN	Q9up79 homo sapien
14	171	46.8	1593	1 AT12_HUMAN	P59397 homo sapien
15	169.5	46.4	1074	1 SMSA_HUMAN	Q13591 homo sapien
16	169	46.3	562	1 AT15_MOUSE	P59384 mus musculus
17	165	45.2	1077	1 AT10_HUMAN	Q9h324 homo sapien
18	165	45.2	1077	1 SMSA_MOUSE	Q62217 mus musculus
19	164	44.9	1170	1 TSP1_BOVIN	Q28178 bos taurus
20	163	44.7	997	1 AT57_HUMAN	Q9ukp4 homo sapien
21	162	44.4	1172	1 TSP2_HUMAN	P35442 homo sapien
22	161	44.1	930	1 AT55_MOUSE	Q9r001 mus musculus
23	161	44.1	1095	1 AT17_HUMAN	Q8te56 homo sapien
24	161	44.1	1170	1 TSP1_HUMAN	P07996 homo sapien
25	161	44.1	1170	1 TSP1_MOUSE	P35441 mus musculus
26	160	43.8	1178	1 TSP2_CHICK	P35440 gallus gall
27	156	42.7	1172	1 TSP2_MOUSE	Q03350 mus musculus
28	155.5	42.6	967	1 AT51_HUMAN	Q9uh18 homo sapien
29	155.5	42.6	967	1 AT51_RAT	Q9wq1 rattus norv
30	155.5	42.6	968	1 AT51_MOUSE	P97857 mus musculus
31	155	42.5	1170	1 TSP2_BOVIN	P59516 bos taurus
32	154.5	42.3	837	1 AT54_HUMAN	Q75173 homo sapien
33	154.5	42.3	860	1 AT56_HUMAN	Q9ukp5 homo sapien

34	153	41.9	930	1 AT55_HUMAN	Q9una0 homo sapien
35	151.5	41.5	630	1 AT54_RAT	Q9esp7 rattus norv
36	150	41.1	1093	1 SMSB_HUMAN	Q9p283 homo sapien
37	150	41.1	1093	1 SMSB_MOUSE	Q60519 mus musculus
38	147	40.3	470	1 PROF_CAVPO	Q64161 cavia porce
39	145.5	39.9	469	1 PROF_HUMAN	P27918 homo sapien
40	143	39.2	1173	1 TSP1_XENLA	P35448 xenopus lae
41	135.5	37.1	437	1 PROF_MOUSE	P11690 mus musculus
42	135	37.0	1572	1 BA12_HUMAN	O60241 homo sapien
43	132	36.2	207	1 AT55_BOVIN	Q9tt92 bos taurus
44	131.5	36.0	843	1 CO7_PIG	Q9tuq3 sus scrofa
45	131	35.9	525	1 AT11_HUMAN	Q8n6g5 homo sapien

ALIGNMENTS

RESULT 1
AT14_HUMAN
ID AT14_HUMAN STANDARD; PRT; 1223 AA.
AC Q8WXS8; Q8TE55; Q8TE58;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-14 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 14) (ADAM-TS 14) (ADAM-TS14).
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Bredelow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
RT metalloproteinase family."
RL Biochim. Biophys. Acta 1522:221-225(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Fetal lung;
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTSs, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains."
RL Gene 283:49-62(2002).
RN [3]
RP SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND
RP ALTERNATIVE PROMOTER USAGE.
RX MEDLINE=21839041; PubMed=11741898;
RA Colige A., Vandenberghe I., Thierry M., Lambert C.A., Van Beeumen J.,
RA Li S.-W., Prockop D.J., Lapiere C.M., Nussgens B.V.;
RT "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying
RT high homology with ADAMTS-2 and ADAMTS-3";
RL J. Biol. Chem. 277:5756-5766(2002).
CC -!- FUNCTION: Has a aminoprocollagen type I activity processing
CC activity in the absence of ADAMTS2. Seems to be synthesized as a
CC latent enzyme that requires activation to display aminoprocollagen
CC peptidase activity.
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (by similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative promoter;
CC Comment=2 isoforms, A (shown here) and B, are produced by use of
CC alternative promoters;
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q8WXS8-1; Sequence=Displayed;
CC Name=B;
CC IsoId=Q8WXS8-2; Sequence=VSP_006958;
CC Name=C;

CC IsoId=Q8WXS9-3; Sequence=VSP_006958, VSP_005501;
CC Notes=Produced by alternative splicing of isoform B;
CC Name=D;
CC IsoId=Q8WXS9-4; Sequence=VSP_005501;
CC Notes=Produced by alternative splicing of isoform A;
CC -!- TISSUE SPECIFICITY: Expressed in retina and at low levels in
CC brain, lung and placenta. High expression in fetal tissues.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF358666; AAL40229.1; -;
CC EMBL; AJ345098; CAC87943.1; -;
CC EMBL; AF362351; AAL79814.1; -;
CC MEROPS; M12.024; -;
CC Genew; HGNC:14899; ADAMTS14.
CC MIM; 607506; -;
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pep_M12B_prosep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp.1; 4.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS00215; ADAM_MEROP; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00900; PLAC; 1.
CC PROSITE; PS00092; TSP1; 4.
CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Collagen degradation; Repeat; Extracellular matrix;
CC Alternative splicing; Alternative promoter usage.
CC SIGNAL 1 22 POTENTIAL.
CC PROPEP 23 252 BY SIMILARITY.
CC CHAIN 253 1223 ADAMTS-14.
CC DOMAIN 253 460 METALLOPROTEASE.
CC DOMAIN 461 551 DISINTEGRIN-LIKE.
CC DOMAIN 552 607 TSP TYPE-1 1.
CC DOMAIN 608 729 CYS-RICH.
CC DOMAIN 730 846 SPACER.
CC DOMAIN 847 907 TSP TYPE-1 2.
CC DOMAIN 908 967 TSP TYPE-1 3.
CC DOMAIN 968 1022 TSP TYPE-1 4.
CC DOMAIN 1059 1097 PLAC.
CC DOMAIN 875 878 POLY-ARG.
CC DOMAIN 1100 1221 PRO-RICH.
CC ACT SITE 399 399 BY SIMILARITY.
CC METAL 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
CC METAL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
CC CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 941 941 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1027 1027 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 1 68 Missing (in isoform B and isoform C).
FT VARSPLIC 368 368 /FTId=VSP_005958:
FT VARSPLIC 368 368 G -> GMQG (in isoform C and isoform D).
FT CONFLICT 868 868 /FTId=VSP_005501:
FT CONFLICT 884 884 Q -> R (IN REF. 2).
FT CONFLICT 901 901 Q -> H (IN REF. 2).
FT CONFLICT 923 923 C -> S (IN REF. 2).
FT CONFLICT 1024 1024 C -> Y (IN REF. 2).
FT CONFLICT 1024 1024 N -> S (IN REF. 2).
SQ SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F170F CRC64;

Query Match 56.4%; Score 206; DB 1; Length 1223;
Best Local Similarity 53.1%; Pred. No. 1.7e-15;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSWGPRSPCRSCGGGWTRRQCNNPRPAFGGRACVGLADLAQMNCNTOACEKTLQLEFM 60
Db 555 WSWTKFGSCRSRSCGGGVRSSRSCNNPSPAYGGRPLGPMFFEQVCNSECPCGVIEDPR 614
QY 61 SQQC 64
Db 615 AQQC 618

RESULT 2
AT16 HUMAN STANDARD; PRT; 1224 AA.
ID AT16 HUMAN
AC Q8TE57; Q8IVE2;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-16 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 16) (ADAM-TS 16) (ADAM-TS16).
GN ADAMTS16 OR KIAA2029.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTS, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
RN [2]
RP SEQUENCE OF 204-1224 FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
CC -!- COPACITOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q8TE57-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q8TE57-2; Sequence=VSP_007664, VSP_007665;
CC -!- TISSUE SPECIFICITY: Expressed in fetal lung and kidney and in
CC adult prostate and ovary.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.


```

1  WSGWPRPSCRS CGGVVTRRQCNPRPAPGGRACVGADLOAEMCNTQACEKTOLEEM 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
589 WSDWSSWSPCSR TC GGGSVSHRSRLCTNPKPSHGKFCGEGSTRTLKCLNSQKCRDSVDFR 648
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61  SQQC 64
   : |||
649 AAQC 652
   |||
db

```

RESULT 3

AT59 HUMAN STANDARD; PRT; 1935 AA.

Q9P2N4; Q9NR29;

16-OCT-2001 (Rel. 40, Created)

10-OCT-2003 (Rel. 42, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)

ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).

ADAMTS9 OR KIAA1312.

Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606;

[1] _RN

SEQUENCE FROM N.A. (ISOFORM 3).

TISSUE=Fetal;

MEDLINE=20396138; PubMed=10936055;

Clark M.E.; Keiner G.S.; Turbeville L.A.; Boyer A.; Arden K.A.; Maki R.A.;

"ADAMTS 9, a novel member of the ADAM-TS/Metallospodin gene family.";

RT

Genomics 67:343-350(2000).

[2]

SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.

MEDLINE=22513925; PubMed=12514189;

Somerville R.P.; Longpre J.-M.; Jungers K.A.; Engle J.M.; Ross M.; Evanko S.; Wight T.N.; Leduc R.; Apté S.S.;

"Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS subfamily related to Caenorhabditis elegans GON-1.";

J. Biol. Chem. 278:9503-9513(2003).

[3]

SEQUENCE OF 159-1935 FROM N.A. (ISOFORM 2).

TISSUE=Brain;

MEDLINE=20181126; PubMed=10718198;

Nagase T.; Kikuno R.; Ishikawa K.-I.; Hirose M.; Ohara O.;

"Prediction of the coding sequences of unidentified human genes. XVI. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro.";

DNA Res. 7:65-73(2000).

CC

-!- FUNCTION: Cleaves the large aggregating proteoglycans, aggrecan and versican.

CC

-!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1838-Glu-|-Ala-1839 site and versican at the 1428-Glu-|-Ala-1429 site.

CC

CC

-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC

-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC

-!- ALTERNATIVE PRODUCTS:

CC

Event=Alternative splicing; Named isoforms=3;

CC

Name=1; Synonyms=ADAMTS-9B;

CC

ISOId=Q9P2N4-3; Sequence=Displayed;

CC

Name=2; Synonyms=Long;

CC

ISOId=Q9P2N4-1; Sequence=VSP_007548, VSP_007549;

CC

Note=May result from the retention of an intron in the cDNA leading to a premature stop codon;

CC

Name=3; Synonyms=Short;

CC

ISOId=Q9P2N4-2; Sequence=VSP_005499, VSP_005500;

CC

-!- TISSUE SPECIFICITY: Highly expressed in all fetal tissues. Expressed in a number of adult tissues with highest expression in heart, placenta and skeletal muscle.

CC

CC

-!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix (By similarity).

CC

CC

```
Query Match      52.6%; Score 192; DB 1; Length 1224;
Best Local Similarity 46.9%; Pred. No. 6.2e-14;
Matches 30: Conservative 12; Mismatches 22; Indels 0; Caps 0;
```

CC -!- DOMAIN: The ancillary domains, including the TSRS domain, are
 CC required for specific extracellular localization and for its
 CC versicanase and aggrecanase activities
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 GON domain.
 CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF261918; AAR89106.1; -
 CC EMBL; AF488803; AA015765.1; -
 CC EMBL; AB037733; BAA92550.1; -
 CC HSSP; P15167; IATL.
 CC DR MEROPS; M12.021; -
 CC DR Genew; HGNC:13202; ADAMTS9.
 CC DR MIM; 605421; -
 CC DR GO; GO:0008237; F:metallopeptidase activity; TAS.
 CC DR GO; GO:0007275; P:development; TAS.
 CC DR GO; GO:0006516; P:glycoprotein catabolism; TAS.
 CC DR InterPro; IPR001762; Disintegrin.
 CC DR InterPro; IPR001818; Pept_M10A_M12B.
 CC DR InterPro; IPR006025; Pept_M_Zn_BS.
 CC DR InterPro; IPR001590; Peptidase M12B.
 CC DR InterPro; IPR002870; Peptidase M12B_N.
 CC DR InterPro; IPR000884; TSP1.
 CC DR Pfam; PF01562; Pep_M12B_propep; 1.
 CC DR Pfam; PF01421; Repolysin; 1.
 CC DR Pfam; PF00090; tsp_1; 11.
 CC DR SMART; SM00209; TSP1; 12.
 CC DR PROSITE; PS02015; ADAM_MEPPO; 1.
 CC DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
 CC DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC DR PROSITE; PS02014; DISINTEGRIN_2; FALSE_NEG.
 CC DR PROSITE; PS00092; TSP1; 14.
 CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 CC KW HydroLase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 CC Repeat; Extracellular matrix; Alternative splicing.
 CC DR SIGNAL 1 18
 CC FT PROPEP 19 287
 CC FT CHAIN 288 1935
 CC FT DOMAIN 293 499
 CC FT DOMAIN 509 587
 CC FT DOMAIN 588 643
 CC FT DOMAIN 644 752
 CC FT DOMAIN 753 877
 CC FT DOMAIN 878 936
 CC FT DOMAIN 939 997
 CC FT DOMAIN 998 1049
 CC FT DOMAIN 1052 1109
 CC FT DOMAIN 1110 1166
 CC FT DOMAIN 1182 1240
 CC FT DOMAIN 1241 1296
 CC FT DOMAIN 1328 1379
 CC FT DOMAIN 1382 1440
 CC FT DOMAIN 1441 1494
 CC FT DOMAIN 1497 1555
 CC FT DOMAIN 1556 1611
 CC FT DOMAIN 1612 1676
 CC FT DOMAIN 1677 1734
 CC FT DOMAIN 1735 1935
 CC FT DOMAIN 88 96
 CC FT SITE 223 223
 CC FT METAL 434 434
 CC FT ACT_SITE 435 435

FT METAL 438 438 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METAL 444 444 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 840 840 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1788 1788 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1806 1806 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1064 1072 CLVTGKGH -> VRWEGCYFP (in isoform 3).
 FT VARSPLIC 1073 1935 Missing (in isoform 3).
 FT VARSPLIC 1624 1629 CSVTG -> VPSNEL (in isoform 2).
 FT VARSPLIC 1630 1935 Missing (in isoform 2).
 FT CONFLICT 46 46 S -> G (IN REF. 1).
 FT CONFLICT 96 96 P -> S (IN REF. 1).
 FT CONFLICT 182 182 D -> G (IN REF. 2).
 FT CONFLICT 367 367 F -> L (IN REF. 1).
 FT CONFLICT 1117 1117 V -> G (IN REF. 3).
 SQ SEQUENCE 1935 AA; 216556 MW; FD3D51E88300A3C6 CRC64;
 Query Match 52.3%; Score 191; DB 1; Length 1935;
 Best Local Similarity 46.9%; Pred. No. 1.2e-13;
 Matches 30; Conservative 11; Mismatches 23; Indels 0; Gaps 0;
 QY 1 WSSWGPRSPSCSGGVWTRRRCNNRPAGGRACVADIQAEWNCNTQACEKTOLEFM 60
 DB 591 WGSWSPFGTCSTCGGKITAIRECNREPKNKGKVCVGRMKFKSCNTEPCLKQKEDFR 650
 QY 61 SQQC 64
 DB 651 DEQC 654
 RESULT 4
 AT18 HUMAN STANDARD; PRT; 1081 AA.
 AC Q8TE60;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE ADAMTS-18 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 18) (ADAM-TS 18) (ADAM-TS18).
 GN ADAMTS18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBITaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21856482; PubMed=11867212;
 RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
 RA Lopez-Otin C.;
 RT "Cloning, expression analysis, and structural characterization of
 RT seven novel human ADAMTSs, a family of metalloproteinases with
 RT disintegrin and thrombospondin-1 domains.";
 RL Gene 283:49-62(2002).
 CC -!- COPACITOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney
 CC and in adult brain, prostate, submaxillary gland, and endothelium.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ311903; CAC83612.1; --
CC Genbank; HGNC:17110; ADAMTS18.
CC MIM; 607512; --
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pept_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 4.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS00215; ADAM_MPRO; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00392; TSP1; 3.
CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 47 POTENTIAL.
FT PROPEP 48 284 BY SIMILARITY.
FT CHAIN 285 1081 ADAMTS-18.
FT DOMAIN 285 497 METALLOPROTEASE.
FT DOMAIN 498 577 DISINTEGRIN-LIKE.
FT DOMAIN 588 643 TSP TYPE-1 1.
FT DOMAIN 644 749 CYS-RICH.
FT DOMAIN 750 876 SPACER.
FT DOMAIN 877 931 TSP TYPE-1 2.
FT DOMAIN 933 991 TSP TYPE-1 3.
FT DOMAIN 993 1048 TSP TYPE-1 4.
FT SITE 254 254 CYSTEINE SWITCH (POTENTIAL).
FT METAL 435 435 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 436 436 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 439 439 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 445 445 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 744 744 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 837 837 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1081 AA; 119655 MW; 0438BF6456764616 CRC64;

Query Match 51.8; Score 189; DB 1; Length 1081;
Best Local Similarity 46.9; Pred. No. 1.2e-13;
Matches 30; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 1 WSSWGRSPSCSGGCVVTRRRONPRPFGGRACVADIQAEWMCNTQACRTQLEPM 60
Db 591 WSAWSKWSKSCRTCCGGVKQERHCNPKPQYGGIFCGSSRIYQLCNINPCNENSLDFR 650

QY 61 SQQC 64
Db 651 AQCQ 654

RESULT 5
AT20 MOUSE
ID AT20 MOUSE
AC P59511; PRT; 1906 AA.
DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-20 precursor (BC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
DE TS20).
GN ADAMTS20.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=Fetal brain;
RC MEDLINE=22562771; PubMed=12562771;
RA Llanazares M., Cal S., Quesada V., Lopez-Otin C.;
RT "Identification and characterization of ADAMTS-20 defines a novel
RT subfamily of metalloproteinases-disintegrins with multiple
RT thrombospondin-1 repeats and a unique GON domain.";
RL J. Biol. Chem. 278:13382-13389(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.
RC STRAIN=DBA/2;
RX MEDLINE=22806432; PubMed=12925592;
RA Rao C., Foerzler D., Loftus S.K., Liu S., McPherson J.D.,
RA Jungers K.A., Apte S.S., Pavan W.J., Beier D.R.;
RT "A defect in a novel ADAMTS family member is the cause of the belted
RT white-spotting mutation.";
RL Development 130:4665-4672(2003).
CC -!- FUNCTION: May play a role in tissue-remodeling process occurring
CC in both normal and pathological conditions.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=ADAMTS20 B long isoform;
CC IsoId=PS9511-1; Sequence=displayed;
CC Name=2; Synonyms=ADAMTS20 A short isoform;
CC IsoId=PS9511-2; Sequence=VSP_007606; VSP_007607;
CC -!- TISSUE SPECIFICITY: Expressed at low level in testis and brain.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- DISEASE: Defects in ADAMTS20 are the cause of the belted (bt)
CC phenotype. It is a pigmental defect which occurs as a result of a
CC defect in melanocyte development.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ512753; CAD54808.3; --
CC EMBL; AY189815; AAC74895.1; --
CC EMBL; AY189816; AAC74896.1; --
CC MGD; MGI:2660628; Adams20.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF01562; Pept_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 11.
CC SMART; SM00209; TSP1; 14.
CC PROSITE; PS00215; ADAM_MPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

DR	PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.	DE	ADAMTS3 OR KIAA0356.	II-NP).
DR	PROSITE; PS50092; TSP1; 13.	GN	Homo sapiens (Human).	
DR	PROSITE; PS50042; ZINC_PROTEASE; 1.	OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
KW	Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;	OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
KW	Repeat; Extracellular matrix; Alternative splicing.	OX	NCBI_TaxID=9606;	
FT	SIGNAL 1 26 POTENTIAL.	EN	[1]	
FT	PROPEP 27 249 BY SIMILARITY.	RP	SEQUENCE OF 1-227 FROM N.A.	
FT	CHAIN 250 1906 ADAMTS-20.	RX	MEDLINE=21402912; PubMed=11408482;	
FT	DOMAIN 250 464 METALLOPROTEASE.	RA	Fernandes R.J., Hirohata S., Engle J.M., Colige A., Cohn D.H.,	
FT	DOMAIN 465 552 DISINTEGRIN-LIKE.	RA	Byre D.R., Apte S.S.;	
FT	DOMAIN 553 608 TSP TYPE-1 1.	RT	"Procollagen II amino propeptide processing by ADAMTS-3. Insights on	
FT	DOMAIN 609 720 CYS-RICH.	RT	dermatoparaxis.";	
FT	DOMAIN 721 842 SPACER.	RL	J. Biol. Chem. 276:31502-31509(2001).	
FT	DOMAIN 843 901 TSP TYPE-1 2.	RL	[2]	
FT	DOMAIN 906 962 TSP TYPE-1 3.	RP	SEQUENCE OF 5-1205 FROM N.A.	
FT	DOMAIN 962 1015 TSP TYPE-1 4.	RC	TISSUE=Brain;	
FT	DOMAIN 1017 1074 TSP TYPE-1 5.	RX	MEDLINE=97349984; PubMed=9205841;	
FT	DOMAIN 1075 1131 TSP TYPE-1 6.	RA	Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,	
FT	DOMAIN 1148 1202 TSP TYPE-1 7.	RA	Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;	
FT	DOMAIN 1203 1260 TSP TYPE-1 8.	RT	"Prediction of the coding sequences of unidentified human genes. VII.	
FT	DOMAIN 1300 1351 TSP TYPE-1 9.	RT	The complete sequences of 100 new cDNA clones from brain which can	
FT	DOMAIN 1354 1411 TSP TYPE-1 10.	RT	code for large proteins in vitro.";	
FT	DOMAIN 1412 1465 TSP TYPE-1 11.	RL	DNA Res. 4:141-150(1997).	
FT	DOMAIN 1468 1526 TSP TYPE-1 12.	CC	-!- FUNCTION: Cleaves the propeptides of type II collagen prior to	
FT	DOMAIN 1527 1584 TSP TYPE-1 13.	CC	fibрил assembly. Does not act on types I and III collagens.	
FT	DOMAIN 1585 1648 TSP TYPE-1 14.	CC	-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).	
FT	DOMAIN 1650 1706 TSP TYPE-1 15.	CC	-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular	
FT	DOMAIN 1707 1906 GON.	CC	matrix (By similarity).	
FT	METAL 399 399 ZINC (CATALYTIC) (BY SIMILARITY).	CC	-!- TISSUE SPECIFICITY: Found in cartilage and skin.	
FT	ACT_SITE 400 400 BY SIMILARITY.	CC	-!- DOMAIN: The spacer domain and the TSP type-1 domains are important	
FT	METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).	CC	for a tight interaction with the extracellular matrix. (By	
FT	METAL 409 409 ZINC (CATALYTIC) (BY SIMILARITY).	CC	-!- PTM: The precursor is cleaved by a furin endopeptidase (By	
FT	CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	similarity).	
FT	CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Belongs to peptidase family M12B.	
FT	CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Contains 1 disintegrin-like domain.	
FT	CARBOHYD 798 798 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Contains 1 PLAC domain.	
FT	CARBOHYD 805 805 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Contains 4 TSP type-1 domains.	
FT	CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- CAUTION: Has sometimes been referred to as ADAMTS4.	
FT	CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-----	
FT	CARBOHYD 1562 1562 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
FT	CARBOHYD 1719 1719 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
FT	CARBOHYD 1759 1759 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	the European Bioinformatics Institute. There are no restrictions on its	
FT	CARBOHYD 1777 1777 N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	use by non-profit institutions as long as its content is in no way	
FT	VARSPPLIC 1424 1425 CS -> VR (in isoform 2).	CC	modified and this statement is not removed. Usage by and for commercial	
FT	VARSPPLIC 1424 1425 /FTid=VSP_007606.	CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
FT	VARSPPLIC 1426 1906 Missing (in isoform 2).	CC	or send an email to license@isb-sib.ch).	
FT	VARSPPLIC 1426 1906 /FTid=VSP_007607.	CC	-----	
FT	CONFLICT 1211 1211 D -> Y (IN REF. 2).	DR	EMBL; AF247668; AAK28400.1; -.	
FT	CONFLICT 1262 1262 S -> L (IN REF. 2).	DR	EMBL; AB002364; BAA20821.1; -.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	MEROPS; M12.220; -.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	Genew; HGNC:219; ADAMTS3.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	MIM; 605011; -.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	GO; GO:0005578; C:extracellular matrix; NAS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	GO; GO:0004222; F:metalloendopeptidase activity; NAS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	GO; GO:0008270; F:zinc ion binding; NAS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	GO; GO:0030574; P:collagen catabolism; NAS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	GO; GO:0030399; P:collagen fibril organization; NAS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR001762; Disintegrin.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR006025; Pept_M_Zn_BS.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR001590; Peptidase M12B.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR002870; Peptidase_M12B_N.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR000884; TSP1.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	InterPro; IPR008085; TSP_i.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	Pfam; PF01562; Pep_M12B_propep; 1.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	Pfam; PF01421; Repolysin; 1.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	Pfam; PF00090; tsp_1; 4.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	PRINTS; PR01705; TSP1REPEAT.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	SMART; SM00209; TSP1; 4.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	PROSITE; PS50215; ADAM_MEPRO; 1.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	PROSITE; PS50214; DISINTEGRIN_2; FALSE_NEG.	
FT	SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;	DR	PROSITE; PS50900; PLAC; 1.	

Query Match	51.5%; Score 188; DB 1; Length 1906;
Best Local Similarity	46.9%; Pred. No. 2.6e-13;
Matches	30; Conservative 10; Mismatches 24; Indels 0; Gaps 0;
QY	1 WSSGPRSPCRSGGGVVTRRRQNNRPAPGGRACVAGDAQAEACNTQACEKTLQEFM 60
Db	556 WGPWGPVSSCSRTCGGKIKSTARLCDPRPNRGGRYCVGRMKFRKSCNTDSCPKGRDPR 615
QY	61 SQQC 64
Db	616 EKQC 619
RESULT 6	
AT53 HUMAN	
ID AT53 HUMAN	STANDARD; PRT; 1205 AA.
AC O15072; Q9BX28;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 28-FEB-2003 (Rel. 41, Last sequence update)	
DT 15-MAR-2004 (Rel. 43, Last annotation update)	
DE ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase	
DE with thrombospondin motifs 3) (ADAM-TS 3) (ADAM-TS3) (Procollagen II	
DE amino-propeptide processing enzyme) (Procollagen II N-proteinase) (PC	

DR PROSITE; P550092; TSP1; 4.
 DR PROSITE; P500142; ZINC PROTEASE; FALSE NEG.
 KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Extracellular matrix; Heparin-binding.
 FT SIGNAL 1 20
 FT PROPEP 21 249 BY SIMILARITY.
 FT CHAIN 250 1205 ADAMTS-3.
 FT DOMAIN 250 469 METALLOPROTEASE.
 FT DOMAIN 470 550 DISINTEGRIN-LIKE.
 FT DOMAIN 551 606 TSP TYPE-1 1.
 FT DOMAIN 608 712 CIS-RICH.
 FT DOMAIN 713 844 SPACER.
 FT DOMAIN 845 905 TSP TYPE-1 2.
 FT DOMAIN 906 965 TSP TYPE-1 3.
 FT DOMAIN 966 1014 TSP TYPE-1 4.
 FT DOMAIN 1015 1054 PLAC.
 FT DOMAIN 246 249 POLY-ARG.
 FT METAL 398 398 ZINC (CATALYTIC) (BY SIMILARITY).
 FT ACT SITE 399 399 BY SIMILARITY.
 FT METL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
 FT METL 408 408 ZINC (CATALYTIC) (BY SIMILARITY).
 FT CARBOHYD 83 83 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1205 AA; 135570 MW; EB07B286FC85FB87 CRC64;
 Query Match 50.7%; Score 185; DB 1; Length 1205;
 Best Local Similarity 48.4%; Pred. No. 3.7e-13;
 Matches 31; Conservative 10; Mismatches 23; Indels 0; Gaps 0;
 QY 1 WSGWGRSPCRSGGGVTRRRRCNNRPAPFGGRACVAGDLQBMCTQACEKQLPEFM 60
 Db 554 WGSWTFGSCRTCTGVRTRTCNNPMPINGGDCGFVFEYQLCTECCQKHFDR 613
 QY 61 SQQC 64
 Db 614 AQQC 617
 RESULT 7
 AT52_BOVIN STANDARD; PRT; 1205 AA.
 AC P79331.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase
 DE with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-TS2) (Procollagen I/II
 DE amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC
 DE I-NP) (Procollagen N-endopeptidase) (pNPI).
 GN ADAMTS2 OR NPI.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin.
 RX MEDLINE=97255960; PubMed=9122202;
 RA Collige A., Li S.W., Sieron A.L., Nuegens B.V., Prockop D.J.,
 RA Lapiere C.M.;
 RT "cDNA cloning and expression of bovine procollagen I N-proteinase: a
 RT new member of the superfamily of zinc-metalloproteinases with binding
 RT sites for cells and other matrix components.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379(1997).
 RN [2]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=95348096; PubMed=7622483;
 RA Collige A., Beschin A., Samyn B., Goebels Y., Van Beeumen J.,
 RA Nuegens B.V., Lapiere C.M.;
 RT "Characterization and partial amino acid sequencing of a 107-kDa
 RT procollagen I N-proteinase purified by affinity chromatography on
 RT immobilized type XIV collagen.";
 RL J. Biol. Chem. 270:16724-16730(1995).
 CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
 CC to fibril assembly. Does not act on type III collagen. May also
 CC play a role in development that is independent of its role in
 CC collagen biosynthesis.
 CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
 CC alpha-1(I) at Pro-[Gln and of alpha-1(II) and alpha-2(I) chains
 CC at Ala-[Gln.
 CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 CC -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to
 CC collagen type XIV.
 CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
 CC matrix (By similarity).
 CC -!- TISSUE SPECIFICITY: Enzymatic activity is detected at high level
 CC in all type I collagen-rich tissues such as skin, bones, tendons
 CC and aorta and at low level in brain and thymus. The mRNA levels
 CC were disproportionately high in heart, liver, retina and muscle.
 CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
 CC for a tight interaction with the extracellular matrix.
 CC -!- PTM: The N-terminus is blocked.
 CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
 CC similarity).
 CC -!- DISEASE: Defects in ADAMTS2 are the cause of dermatosparaxis, a
 CC recessively inherited disorder characterized by severe skin
 CC fragility and biochemically by the presence in skin of procollagen
 CC incompletely processed at the N-terminus.
 CC -!- SIMILARITY: Belongs to peptidase family M12B.
 CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
 CC -!- SIMILARITY: Contains 1 PLAC domain.
 CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
 CC -!- CAUTION: Has sometimes been referred to as ADAMTS3.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC -----
 CC EMBL; X96389; CAA65253.1; -.
 CC FIR; T18517; T18517.
 CC MEROPS; M12.301; -.
 CC InterPro; IPR001762; Disintegrin.
 CC InterPro; IPR006025; Pept_M_Zn_BS.
 CC InterPro; IPR001590; Peptidase_M12B.
 CC InterPro; IPR002870; Peptidase_M12B_N.
 CC InterPro; IPR000884; TSP1.
 CC Pfam; PF01562; Pep_M12B_propep; 1.
 CC Pfam; PF01421; Reprolysin; 1.
 CC Pfam; PF00090; tsp.1; 4.
 CC SMART; SM00209; TSP1; 4.
 CC PROSITE; PS0215; ADAM_MEPRO; 1.
 CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 CC PROSITE; PS00900; PLAC; 1.
 CC PROSITE; PS00092; TSP1; 4.
 CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
 KW Repeat; Collagen degradation; Extracellular matrix.
 FT SIGNAL 1 28
 FT PROPEP 29 253 BY SIMILARITY.
 FT CHAIN 254 1205 ADAMTS-2.
 FT DOMAIN 254 473 METALLOPROTEASE.
 FT DOMAIN 474 554 DISINTEGRIN-LIKE.
 FT DOMAIN 555 610 TSP TYPE-1 1.
 FT DOMAIN 612 716 CIS-RICH.
 FT DOMAIN 717 845 SPACER.

```
FT DOMAIN 848 906 TSP TYPE-1 2.
FT DOMAIN 908 968 TSP TYPE-1 3.
FT DOMAIN 969 1023 TSP TYPE-1 4.
FT DOMAIN 1053 1091 PLAC.
FT DOMAIN 31 35 POLY-ALA.
FT DOMAIN 177 180 POLY-GLU.
FT METAL 402 403 ZINC (CATALYTIC) (POTENTIAL).
FT ACT_SITE 403 403 BY SIMILARITY.
FT METAL 406 406 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).
FT SITE 685 687 CELL ATTACHMENT SITE (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 245 245 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 942 942 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 943 943 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 987 987 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1025 1025 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1092 1092 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1139 1139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1144 1144 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1205 AA; 133887 MW; 7B5B232A45320371 CRC64;

Query Match 50.1%; Score 183; DB 1; Length 1205;
Best Local Similarity 46.9%; Pred. No. 6.2e-13;
Matches 30; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY 1 WSGVPRSPGSRSCGGVTRRRQNNPRPAFGGRACVADLQAEKNTQACEKTLQLEPM 60
DB 558 WGAWSPPGSRCTCGTGVKTRQCDNPHNGGRTCSGLAYDFQLCNSQDCPDALADPR 617

QY 61 SQQC 64
DB 618 EEQC 621

RESULT 8
ID AT52_MOUSE STANDARD; PRT; 1213 AA.
AC Q8C9W3;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 2) (ADAM-TS 2) (ADAM-17) (Procollagen I/II
DE amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC
DE I-NP) (Procollagen N-endopeptidase) (pNP).
GN ADAMTS2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22354683; PubMed=12466851;
RA Okazaki Y., Furuto M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nikaide I., Opat N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojohori T.,
RA Badaricelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kapapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.B., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gliss C., Godzik A., Gough J.,
RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sautelina A., Schneider C., Semple C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Vezardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
```

```
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N., K.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,
RA Shiraki T., Waki K., Kawai K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
[2]
RP SEQUENCE FROM N.A.
RX STRAIN=EVB/N; TISSUE=Breast tumor;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.P., Zesberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Willalson D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Brakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA Schnarch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
CC -I- FUNCTION: Cleaves the propeptides of type I and II collagen prior
CC to fibril assembly. Does not act on type III collagen. May also
CC play a role in development that is independent of its role in
CC collagen biosynthesis (By similarity).
CC -I- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
CC alpha-1(I) at Pro-|-Gln and of alpha-1(II) and alpha-2(I) chains
CC at Ala-|-Gln.
CC -I- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -I- SUBUNIT: May belong to a multimeric complex. Binds specifically to
CC collagen type XIV (By similarity).
CC -I- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -I- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -I- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -I- SIMILARITY: Belongs to peptidase family M12B.
CC -I- SIMILARITY: Contains 1 disintegrin-like domain.
CC -I- SIMILARITY: Contains 1 PLAC domain.
CC -I- SIMILARITY: Contains 4 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AK040370; AAC30572.1; -
CC EMBL; BC046456; BAH46456.1; -
CC MGD; MGI:1347356; Adamts2.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
```

DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tspl1; 4.
DR SMART; SM00209; TSPI; 4.
DR PROSITE; PS02015; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS02014; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00900; PLAC; 1.
DR PROSITE; PS00092; TSPI; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Collagen degradation; Extracellular matrix.
FT SIGNAL 1 28
FT PROPEP 29 260
FT CHAIN 29 1213
FT DOMAIN 261 471
FT DOMAIN 480 560
FT DOMAIN 561 617
FT DOMAIN 618 722
FT DOMAIN 723 851
FT DOMAIN 852 913
FT DOMAIN 914 975
FT DOMAIN 976 1030
FT DOMAIN 1031 1098
FT SITE 692 694
FT METAL 409 409
FT ACT SITE 410 410
FT METAL 413 413
FT METAL 419 419
FT CARBOHYD 111 111
FT CARBOHYD 252 252
FT CARBOHYD 949 949
FT CARBOHYD 950 950
FT CARBOHYD 994 994
FT CARBOHYD 1032 1032
FT CARBOHYD 1099 1099
FT CARBOHYD 1147 1147
FT CARBOHYD 1152 1152
FT CARBOHYD 1152 1152
FT CARBOHYD 1213 1213
FT SEQUENCE 1213 AA; 135298 MW; B27431E00443EDB5 CRC64;
Query Match 49.6%; Score 181; DB 1; Length 1213;
Best Local Similarity 46.9%; Pred. No. 1.1e-12;
Matches 30; Conservative 8; Mismatches 26; Indels 0; Gaps 0;
QY 1 WSSWPRSPSCSGGVVTRRQCNPNPRAFGACVADLQEMCNQACEKTLQLEPM 60
DB 565 WGAWTFPGSCRTGTVKFRTRQCNPNPANGRTCSGLAYDQLGNCNPQCPNSLADFR 624
QY 61 SQQC 64
DB 625 EQQC 628
RESULT 9
AT20 HUMAN STANDARD; PRT; 1911 AA.
AC P59510;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-20 precursor (EC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
DE TS20).
GN ADAMTS20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wight T.N., Leduc R., Apté S.S.;
RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS

subfamily related to Caenorhabditis elegans GON-1.";
J. Biol. Chem. 278:9503-9513(2003).
[2]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
TISSUE=Liver;
MEDLINE=22566039; PubMed=12562771;
Llamazares M., Cal S., Quesada V., Lopez-Otin C.;
"Identification and characterization of ADAMTS-20 defines a novel
subfamily of metalloproteinases-disintegrins with multiple
thrombospondin-1 repeats and a unique GON domain.";
J. Biol. Chem. 278:13382-13389(2003).
CC -!- FUNCTION: May play a role in tissue-remodeling process occurring
in both normal and pathological conditions.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59510-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59510-2; Sequence=VSP_007106, VSP_007107, VSP_007108;
CC -!- TISSUE SPECIFICITY: Very sparingly expressed, although is detected
at low levels in testis, prostate, ovary, heart, placenta, lung
and pancreas. Overexpressed in several brain, colon and breast
carcinomas.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; AF488804; AAC15766.1; -.
EMBL; AJ515153; CAD56159.3; -.
EMBL; AJ515154; CAD56160.2; -.
Genew; HGNC:17178; ADAMTS20.
InterPro; IPR001762; Disintegrin.
InterPro; IPR006025; Pept M Zn BS.
InterPro; IPR001590; Peptidase_M12B.
InterPro; IPR002870; Peptidase_M12B_N.
InterPro; IPR000884; TSPI.
Pfam; PF01562; Pep_M12B_propep; 1.
Pfam; PF01421; Reprolysin; 1.
Pfam; PF00090; tspl1; 11.
SMART; SM00209; TSPI; 12.
PROSITE; PS02015; ADAM_MEPRO; 1. FALSE_NEG.
PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
PROSITE; PS02014; DISINTEGRIN_2; FALSE_NEG.
PROSITE; PS00092; TSPI; 12.
PROSITE; PS00142; ZINC_PROTEASE; 1.
Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
Repeat; Extracellular matrix; Alternative splicing.
FT SIGNAL 1 21
FT PROPEP 22 253
FT CHAIN 254 1911
FT DOMAIN 254 467
FT DOMAIN 468 555
FT DOMAIN 556 611
FT DOMAIN 612 723
FT DOMAIN 724 846
FT DOMAIN 847 905
FT DOMAIN 906 962
FT DOMAIN 967 1024
FT DOMAIN 1025 1074

FT	DOMAIN	1077	1136	TSP TYPE-1 6.	DT	28-FEB-2003 (Rel. 41, Last sequence update)
FT	DOMAIN	1153	1207	TSP TYPE-1 7.	DT	28-FEB-2003 (Rel. 41, Last annotation update)
FT	DOMAIN	1208	1265	TSP TYPE-1 8.	DE	ADAMTS-15 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 15) (ADAM-TS 15) (ADAM-TS15).
FT	DOMAIN	1305	1357	TSP TYPE-1 9.	GN	ADAMTS15.
FT	DOMAIN	1359	1417	TSP TYPE-1 10.	OS	Homo sapiens (Human).
FT	DOMAIN	1418	1476	TSP TYPE-1 11.	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
FT	DOMAIN	1477	1532	TSP TYPE-1 12.	OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
FT	DOMAIN	1536	1589	TSP TYPE-1 13.	OX	NCBI_TaxID=9606;
FT	DOMAIN	1590	1633	TSP TYPE-1 14.	RN	[1]
FT	DOMAIN	1655	1711	TSP TYPE-1 15.	RP	SEQUENCE FROM N.A.
FT	DOMAIN	1712	1911	GON.	RX	MEDLINE=21856482; PubMed=11867212;
FT	METAL	403	403	ZINC (CATALYTIC) (BY SIMILARITY).	RA	Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
FT	ACT SITE	404	404	BY SIMILARITY.	RA	Lopez-Otin C.;
FT	METAL	407	407	ZINC (CATALYTIC) (BY SIMILARITY).	RT	"Cloning, expression analysis, and structural characterization of
FT	METAL	413	413	ZINC (CATALYTIC) (BY SIMILARITY).	RT	seven novel human ADAMTSs, a family of metalloproteinases with
FT	CARBOHYD	92	92	N-LINKED (GLCNAC. . .) (POTENTIAL).	RL	disintegrin and thrombospondin-1 domains.";
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	Gene 283:49-62(2002).
FT	CARBOHYD	445	445	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- COPACTOR: Binds 1 zinc ion per subunit (By similarity).
FT	CARBOHYD	702	702	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
FT	CARBOHYD	717	717	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	matrix (By similarity).
FT	CARBOHYD	728	728	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- TISSUE SPECIFICITY: Expressed in fetal liver and kidney, but not
FT	CARBOHYD	809	809	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	in any of the adult tissues examined
FT	CARBOHYD	870	870	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- DOMAIN: The spacer domain and the TSP type-1 domains are important
FT	CARBOHYD	1062	1062	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	for a tight interaction with the extracellular matrix (By
FT	CARBOHYD	1457	1457	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	similarity).
FT	CARBOHYD	1543	1543	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- PTM: The precursor is cleaved by a furin endopeptidase (By
FT	CARBOHYD	1573	1573	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	similarity).
FT	CARBOHYD	1764	1764	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Belongs to peptidase family M12B.
FT	CARBOHYD	1782	1782	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Contains 1 disintegrin-like domain.
FT	CARBOHYD	1853	1853	N-LINKED (GLCNAC. . .) (POTENTIAL).	CC	-!- SIMILARITY: Contains 3 TSP type-1 domains.
FT	VARSP LIC	683	693	THDLCVQGCQM -> SYNIDNCVLK (in isoform 2).	CC	-----
FT				/FTId=VSP_007106.	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
FT				CSACGGRKRVFCIDQFQKLEDTNCSQVQKPEPHKAC	CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
FT				RSVRCPSWKANSWNEVTSVCGVQQRDVYCRLK -> EDL	CC	the European Bioinformatics Institute. There are no restrictions on its
FT				KYLLPQRTILWELMKNIFCHGKSHMYLINVVDHLLYP	CC	use by non-profit institutions as long as its content is in no way
FT				RHCDPEITETVFLSLWSLQFTWGLDKYKNSL (in	CC	modified and this statement is not removed. Usage by and for commercial
FT				isoform 2).	CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/
FT				/FTId=VSP_007107.	CC	or send an email to license@isb-sib.ch).
FT				Missing (in isoform 2).	CC	-----
FT				/FTId=VSP_007108.	DR	EMBL; AJ315733; CAC86014.1; -
FT				T -> V (IN REF. 2).	DR	MEROPS; M12.025; -
FT				E -> Y (IN REF. 2).	DR	Genew; HGNC:16305; ADAMTS15.
FT				T -> M (IN REF. 2).	DR	MIM; 607509; -
FT				I -> E (IN REF. 2).	DR	InterPro; IPR001762; Disintegrin.
FT				D -> V (IN REF. 2).	DR	InterPro; IPR001818; Pept M10A_M12B.
FT				MISSING (IN REF. 2).	DR	InterPro; IPR006025; Pept M_Zn_BS.
FT				LILQ -> ILIE (IN REF. 2).	DR	InterPro; IPR001590; Peptidase M12B.
FT				GMLAK -> WHVIG (IN REF. 2).	DR	InterPro; IPR002870; Peptidase_M12B_N.
FT				S -> Q (IN REF. 2).	DR	InterPro; IPR000884; TSP1.
FT				T -> R (IN REF. 2).	DR	Pfam; PF01562; Pep_M12B_propep; 1.
FT				T -> A (IN REF. 2).	DR	Pfam; PF01421; Repolysin; 1.
FT				E -> Q (IN REF. 2); CAD56159).	DR	Pfam; PF00090; tsp.1; 3.
FT				1911 AA; 214656 MW; CF592E220D32B250 CRC64;	DR	SMART; SM00209; TSP1; 3.
FT				Query Match 49.6%; Score 181; DB 1; Length 1911;	DR	PROSITE; PS00215; ADAM_MEROP; 1.
FT				Best Local Similarity 45.3%; Pred No. 1,5e-12;	DR	PROSITE; PS00546; CYSTEINE_SWITCH; FALSE NEG.
FT				Matches 29; Conservative 9; Mismatches 26; Indels 0; Gaps 0;	DR	PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
Qy	1 WSSWGPSPSCSCGGVYVTRRQCNRPAPFGGRACVGADLQAEWCNTQACEKTKOLEFM 60				DR	PROSITE; PS00214; DISINTEGRIN_2; FALSE NEG.
Db	559 WGPWEPYSSCSRTCGGIESAIFRCNRNRPBNGNVCVRMKFRSCNTDSCPKGTQDFR 618				DR	PROSITE; PS00092; TSP1; 3.
Qy	61 SQQC 64				DR	PROSITE; PS00143; ZINC_PROTEASE; 1.
Db	619 EKQC 622				KW	Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
					KW	Repeat; Extracellular matrix.
					FT	SIGNAL 1 17 POTENTIAL.
					FT	PROPEP 18 212 BY SIMILARITY.
					FT	CHAIN 213 950 ADAMTS-15.
					FT	DOMAIN 213 427 METALLOPROTEASE.
					FT	DOMAIN 428 515 DISINTEGRIN-LIKE.
					FT	DOMAIN 516 571 TSP TYPE-1 1.
					FT	DOMAIN 572 700 CYS-RICH.
					FT	DOMAIN 701 838 SPACER.
					FT	DOMAIN 839 895 TSP TYPE-1 2.
					FT	DOMAIN 896 949 TSP TYPE-1 3.
					FT	DOMAIN 174 174 CYSTEINE SWITCH (POTENTIAL).
					FT	SITE
RESULT 10						
AT15_HUMAN						
ID AT15 HUMAN						
AC Q8TE58;						
DT 28-FEB-2003 (Rel. 41, Created)						


```
FT CARBOHYD 1145 1145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIT 544 566 HCFKHCIMLTDPILKRDGSMGA -> PRFGAVAHACYPST
FT LGGQGRWIA (in isoform SPNP1).
FT FTID-VSP 005497.
FT VARSPLIT 567 1211 Missing (in isoform SPNP1).
FT /FTID-VSP_005498.
SQ SEQUENCE 1211 AA; 134722 MW; BCSEEF25C23CAD2D CRC64;

Query Match 48.5%; Score 177; DB 1; Length 1211;
Best Local Similarity 45.3%; Pred. No. 2.9e-12;
Matches 29; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 1 WSSWGPRSPCSRSCGGVTRRRQCNPRPAGFGRACVAGDLQAEWMCNTQACEKTOLEFM 60
DB 564 WGAWSFFGSCRTCGTGVKFTQCDNPHFANGRTCSGLAYDQLQSCQPCDPSLADFR 623

QY 61 SQQC 64
DB 624 EQQC 627

RESULT 12
AT88_MOUSE
ID AT88_MOUSE STANDARD; PRT; 905 AA.
AC P57110;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).
GN ADAMTS8.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20079168; PubMed=10610729;
RA Georgiadis K.E., Hirohata S., Seidlin M.F., Apte S.S.;
RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
RL mouse chromosome 9 and human chromosome 11.";
RL Genomics 62:312-315(1999).
CC -!- FUNCTION: Has anti-angiogenic properties (By similarity).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart
CC and low expression during mouse development.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF175282; AAF25805.1; --
CC HSSP; P34179; 1IAG.
CC MEROPS; M12.226; --
CC MGD; MGI:11353468; Adamts8.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR000884; TSP1.
```

```
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; TSP_1; 2.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS0215; ADAM MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE NEG.
DR PROSITE; PS00092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 28 POTENTIAL.
FT PROPEP 29 228 BY SIMILARITY.
FT CHAIN 229 905 ADAMTS-8.
FT DOMAIN 229 452 METALLOPROTEASE.
FT DOMAIN 453 541 DISINTEGRIN-LIKE.
FT DOMAIN 542 597 TSP TYPE-1 1.
FT DOMAIN 599 705 CYS-RICH.
FT DOMAIN 706 847 SPACER.
FT DOMAIN 848 904 TSP TYPE-1 2.
FT METAL 378 378 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 379 379 BY SIMILARITY.
FT METAL 382 382 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 480 480 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 506 506 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 905 AA; 98879 MW; 124D4132933A0CAE CRC64;

Query Match 47.9%; Score 175; DB 1; Length 905;
Best Local Similarity 43.8%; Pred. No. 3.8e-12;
Matches 28; Conservative 9; Mismatches 27; Indels 0; Gaps 0;

QY 1 WSSWGPRSPCSRSCGGVTRRRQCNPRPAGFGRACVAGDLQAEWMCNTQACEKTOLEFM 60
DB 545 WGPWFPGQCSRTCGGCIQFSNRCDNPMQNGRFLGVRKYQSCNTECPNGKSPR 604

QY 61 SQQC 64
DB 605 EQQC 608

RESULT 13
AT88_HUMAN
ID AT88_HUMAN STANDARD; PRT; 890 AA.
AC Q9UP79; Q9NZS0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)
DE (METH-8).
GN ADAMTS8 OR METH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=99367466; PubMed=10438512;
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
RA Lombardo M., Iruela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT family of proteins with angio-inhibitory activity.";
RL J. Biol. Chem. 274:23349-23357(1999).
RN [2]
RP SEQUENCE OF 195-440 FROM N.A.
RX MEDLINE=20079168; PubMed=10610729;
RA Georgiadis K.E., Hirohata S., Seidlin M.F., Apte S.S.;
RA "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
```

```
RT mouse chromosome 9 and human chromosome 11." ;
RL Genomics 62:312-315(1999).
CC -!- FUNCTION: Has anti-angiogenic properties.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower
CC expression in brain, placenta, heart, stomach and fetal brain and
CC kidney.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF060153; AAD48081.1; -.
CC EMBL; AF152883; AAF25806.1; -.
CC HSSP; P34179; 1IAG.
CC MEROPS; M12.226; -.
CC Genew; HGNC:224; ADAMTS8.
CC MIM; 605175; -.
CC GO; GO:0003178; F.integrin binding; TAS.
CC GO; GO:0008237; F.metalloproteinase activity; TAS.
CC GO; GO:0007345; P.embryogenesis and morphogenesis; TAS.
CC GO; GO:0008285; P.negative regulation of cell proliferation; TAS.
CC InterPro; IPR001762; Disintegrin
CC InterPro; IPR006025; Pept_M12B.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pfam_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.
CC Pfam; PF00090; tsp_1; 2.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC PROSITE; PS00215; ADAM_MEROP; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS00092; TSP1; 2.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 27
FT PROPEP 28 214
FT CHAIN 215 890
FT DOMAIN 215 430
FT DOMAIN 439 526
FT DOMAIN 527 582
FT DOMAIN 584 690
FT DOMAIN 691 832
FT DOMAIN 834 889
FT DOMAIN 202 205
FT METAL 364 364
FT ACT_SITE 365 365
FT METAL 368 368
FT METAL 374 374
FT CARBOHYD 345 345
FT CARBOHYD 401 401
FT CARBOHYD 466 466
FT CARBOHYD 491 491
FT CARBOHYD 600 600
FT CONFLICT 195 195
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 27
FT PROPEP 28 214
FT CHAIN 215 890
FT DOMAIN 215 430
FT DOMAIN 439 526
FT DOMAIN 527 582
FT DOMAIN 584 690
FT DOMAIN 691 832
FT DOMAIN 834 889
FT DOMAIN 202 205
FT METAL 364 364
FT ACT_SITE 365 365
FT METAL 368 368
FT METAL 374 374
FT CARBOHYD 345 345
FT CARBOHYD 401 401
FT CARBOHYD 466 466
FT CARBOHYD 491 491
FT CARBOHYD 600 600
FT CONFLICT 195 195
E -> R (IN REF. 2).
FT CONFLICT 413 440 YLTLELDGGHGDCLLDAPGAALPLPTGL -> FSCCHLQGW
FT INPKYLCKVSELKCDLMP (IN REF. 2).
SQ SEQUENCE 890 AA; 96671 MW; 57D70EE03D5739D3 CRC64;
Query Match 47.1%; Score 172; DB 1; Length 890;
Best Local Similarity 43.8%; Pred. No. 8e-12; Indels 0; Gaps 0;
Matches 28; Conservative 9; Mismatches 27;
QY 1 WSGWSPRSCSRSCGGWVTRRRQCNNPRPAFGGRACVGDLOAEMCNQACEKTLQLEPM 60
Db 530 WAPWPGWGCSTCGGVQFSPHRECKDPQNGRYCLGRRAKYQSCHTBEPDGKSPR 589
QY 61 SQQC 64
Db 590 EQQC 593
RESULT 14
AT12 HUMAN
ID AT12 HUMAN STANDARD; PRT; 1593 AA.
AC P58397;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-12 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
GN ADAMTS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RX MEDLINE=21264577; PubMed=11279086;
RA Cal S., Argueelles J.M., Fernandez P.L., Lopez-Otin C.;
RT "Identification, characterization, and intracellular processing of
RT ADAM-TS12, a novel human disintegrin with a complex structural
RT organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940(2001).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC expressed in gastric carcinomas and in cancer cells of diverse
CC origin.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase.
CC -!- PTM: IS SUBJECTED TO AN INTRACELLULAR MATURATION PROCESS LEADING
CC TO A FRAGMENT CONTAINING THE N-TERMINAL REGION INCLUDING THE
CC METALLOPROTEINASE, DISINTEGRIN-LIKE, CYS-RICH AND TS-1 DOMAINS AND
CC THE C-TERMINAL FRAGMENT CONTAINING THE SPACER 2 AND THE FOUR TS-1
CC DOMAINS.
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 8 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AJ250725; CAC20419.1; -.
CC Genew; HGNC:14605; ADAMTS12.
CC MIM; 606184; -.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M10A_M12B.
```

DR InterPro; IPR006025; Pept_M_zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP1.
DR Pfam; PF01562; Pept_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 8.
DR PROSITE; PS0215; ADAM_MPRO; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0900; PLAC; 1.
DR PROSITE; PS0092; TSP1; 6.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 240 BY SIMILARITY.
FT CHAIN 241 1593 ADAMTS-12
FT DOMAIN 241 464 METALLOPROTEASE.
FT DOMAIN 465 544 DISINTEGRIN-LIKE.
FT DOMAIN 542 597 TSP TYPE-1 1.
FT DOMAIN 597 700 CYS-RICH.
FT DOMAIN 701 826 SPACER 1.
FT DOMAIN 823 882 TSP TYPE-1 2.
FT DOMAIN 886 942 TSP TYPE-1 3.
FT DOMAIN 943 996 TSP TYPE-1 4.
FT DOMAIN 996 1315 SPACER 2.
FT DOMAIN 1312 1365 TSP TYPE-1 5.
FT DOMAIN 1367 1421 TSP TYPE-1 6.
FT DOMAIN 1422 1470 TSP TYPE-1 7.
FT DOMAIN 1471 1531 TSP TYPE-1 8.
FT DOMAIN 1534 1574 PLAC.
FT DOMAIN 302 305 POLY-GLU.
FT SITE 208 208 CYSTEINE SWITCH (POTENTIAL).
FT METAL 392 392 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 393 393 BY SIMILARITY.
FT METAL 396 396 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 402 402 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 215 215 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 485 485 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 685 685 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1300 1300 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1371 1371 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1503 1503 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1593 AA; 177545 MW; 07F9F4BE63BD83A3 CRC64;
Query Match 46.8%; Score 171; DB 1; Length 1593;
Best Local Similarity 43.8%; Pred. No. 1.8e-11;
Matches 28; Conservative 5; Mismatches 31; Indels 0; Gaps 0;
QY 1 WSWGPRSPGSRSCGGVTVRRQNNPRPFGGACVGDIAQSMCNQACEKTLQEFM 60
Db 545 WGRSPWHSCHSRTGAGVQSARLNCNPNPEFGGKYCTGERKRYLCLNVHPCRSEAPTR 604
QY 61 SQQC 64
Db 605 QMQC 608
RESULT 15
SMSA_HUMAN

SM5A_HUMAN STANDARD; PRT; 1074 AA.
Q13591; O60408;
30-MAY-2000 (Rel. 39, Created)
30-MAY-2000 (Rel. 39, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 5A precursor (Semaphorin F) (Sema F).
CN SEMA5A OR SEMAF.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RX MEDLINE=98125554; PubMed=9464278;
RA Simmons A.D.; Puschel A.W.; McPherson J.D.; Overhauser J.; Lovett M.;
RT "Molecular cloning and mapping of human semaphorin F from the Cri-du-
chat candidate interval.";
RL Biochem. Biophys. Res. Commun. 242:685-691(1998).
RN [2]
RP SEQUENCE OF 1-494 FROM N.A.
RA Kalicki J.; Harmon G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
DR EMBL; U52840; AAC09473.1; -.
DR EMBL; AC004615; AAC14668.1; -.
DR PIR; JC5928; JC5928.
DR Genew; HGNC:10736; SEMA5A.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR GO; GO:0007267; P:cell-cell signaling; TAS.
DR GO; GO:0007399; P:neurogenesis; TAS.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002185; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; tsp_1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS00092; TSP1; 6.
KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 1074 SEMAPHORIN 5A.
FT DOMAIN 23 968 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 969 989 POTENTIAL.
FT DOMAIN 990 1074 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 226 507 Sema.
FT DOMAIN 540 593 TSP TYPE-1 1.
FT DOMAIN 595 651 TSP TYPE-1 2.
FT DOMAIN 653 702 TSP TYPE-1 3.
FT DOMAIN 707 765 TSP TYPE-1 4.
FT DOMAIN 784 839 TSP TYPE-1 5.
FT DOMAIN 841 896 TSP TYPE-1 6.
FT DOMAIN 897 944 TSP TYPE-1 7.
FT CARBOHYD 142 168 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD	227	227	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	277	277	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	323	323	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	367	367	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	437	437	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	536	536	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	591	591	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	717	717	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CARBOHYD	933	933	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT CONFLICT	56	56	A -> V (IN REF. 2).	
FT CONFLICT	149	149	A -> T (IN REF. 2).	
FT CONFLICT	382	382	V -> M (IN REF. 2).	
FT CONFLICT	494	494	S -> R (IN REF. 2).	
SQ SEQUENCE	1074	AA; 120570 MW; EE3DB763CBE29407 CRC64;		

Query Match 46.4%; Score 169.5; DB 1; Length 1074;
Best Local Similarity 37.7%; Pred. No. 1.8e-11;
Matches 29; Conservative 12; Mismatches 23; Indels 13; Gaps 1;

QY	1	WSWGPRSPCSRCSGGVVTTRQCNPRPAFGGRACVGADLQAEPMCNTQ-----	50
Db	844	WSCWSPWTKCSATCGGHHYMRTRSCSNPAPAYGGDCLGLHTEALCNTQCPESWSEWS	903
QY	51	---ACEKTQLEFMSQOC 64	
Db	904	DWSECEASGVQVRARQC 920	

Search completed: March 13, 2004, 07:39:57
Job time : 2.33333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 6.33333 Seconds
(without alignments)

3188.394 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473

Perfect score: 365

Sequence: 1 WSSWGRSPCSRSCGGVVT.....EMCNTQACEKTQLEFMSQQC 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25.*

1: sp_archaea.*

2: sp_bacteria.*

3: sp_fungi.*

4: sp_human.*

5: sp_invertebrate.*

6: sp_mammal.*

7: sp_mhc.*

8: sp_organelle.*

9: sp_phase.*

10: sp_plant.*

11: sp_todent.*

12: sp_virus.*

13: sp_vertebrate.*

14: sp_unclassified.*

15: sp_rvirus.*

16: sp_bacteriap.*

17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	365	100.0	364	Q9UGQ1	Q9ugq1 homo sapien
2	365	100.0	1427	Q96L37	Q96l37 homo sapien
3	197	54.0	900	Q8K206	Q8k206 mus musculu
4	187	51.2	1092	Q8BKAI	Q8bkai mus musculu
5	173	47.4	769	Q8XRL5	Q8xrl5 drosophila
6	173	47.4	1059	Q9W493	Q9w493 drosophila
7	172	47.1	2673	Q4G6SC3	Q4g6sc3 homo sapien
8	172	47.1	5636	Q4G6RW7	Q4g6rw7 homo sapien
9	170	46.6	1070	Q8CG28	Q8cg28 mus musculu
10	168	46.0	1688	Q8XB0	Q8xb0 drosophila
11	167	45.8	1009	Q8BK1	Q8bk1 mus musculu
12	167	45.8	1800	Q81IB3	Q81ib3 mus musculu
13	161	44.1	269	Q9GL54	Q9gl54 oryctolagus
14	161	44.1	623	Q8BGP4	Q8bgp4 mus musculu
15	161	44.1	1081	Q9U631	Q9u631 drosophila
16	161	44.1	1083	Q9VTT0	Q9vtt0 drosophila

17	161	44.1	1091	5	Q7YU67	Q7yu67 drosophila
18	161	44.1	1171	11	Q8CGB2	Q8cgb2 mus musculu
19	161	44.1	1171	11	Q8CGB2	Q8cgb2 mus musculu
20	159	43.6	89	11	Q9R150	Q9r150 cavia porce
21	159	43.6	921	5	Q969A3	Q969a3 brachiosto
22	157	43.0	741	4	Q96K89	Q96k89 homo sapien
23	156	42.7	1172	11	Q8CG21	Q8cg21 mus musculu
24	156	42.7	1172	11	Q8CG21	Q8cg21 mus musculu
25	155.5	42.6	759	6	Q8H2M8	Q8h2m8 equus cabal
26	155.5	42.6	967	4	Q8NE26	Q8ne26 homo sapien
27	154.5	42.3	833	11	Q8K384	Q8k384 mus musculu
28	154.5	42.3	845	11	Q8BNJ2	Q8bnj2 mus musculu
29	154	42.2	335	11	Q8BVQ2	Q8bvq2 mus musculu
30	154	42.2	844	11	Q8BXU8	Q8bxu8 mus musculu
31	153.5	42.1	1280	11	Q8EPX2	Q8epx2 mus musculu
32	152.5	41.8	839	6	Q7YS95	Q7ys95 bos taurus
33	152	41.6	1637	6	Q9XSV8	Q9xsv8 bos taurus
34	152	41.6	5146	6	Q8SPM4	Q8spm4 bos taurus
35	150.5	41.2	2165	5	Q19791	Q19791 caenorhabdi
36	150	41.1	478	11	Q8BVE5	Q8bve5 mus musculu
37	150	41.1	1122	11	Q7TT33	Q7tt33 mus musculu
38	150	41.1	4998	11	Q8CG65	Q8cg65 mus musculu
39	146.5	40.1	129	4	Q8IVU0	Q8ivu0 homo sapien
40	146.5	40.1	1235	4	Q95428	Q95428 homo sapien
41	146	40.0	957	11	Q7TSK7	Q7tsk7 mus musculu
42	145.5	39.9	3198	5	Q9UG8	Q9ug8 marduca sex
43	144.5	39.6	1461	5	Q8MYA8	Q8mya8 caenorhabdi
44	142	38.9	4123	4	Q75851	Q75851 homo sapien
45	140.5	38.5	1487	5	Q8MPV5	Q8mpv5 caenorhabdi

ALIGNMENTS

RESULT 1

Q9UGQ1 PRELIMINARY; PRT; 364 AA.

AC Q9UGQ1;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN C9ORF8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Young J.M., Woodward K.J., Aziz S., Burley M., Kwiatkowski D.J.,
RA Povey S.;
RT "Cloning of a sugar transporter gene, a G-beta subunit like gene and
RT three novel genes in human chromosome 9q34.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ011374; CAB66157.1; --
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF00090; tsp_1; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 1.
DR PROSITE; PS50092; TSP1; 1.
KW Hypothetical protein.
SQ SEQUENCE 364 AA; 39864 MW; DA42FC5F5345F3A0 CRC64;

Query Match 100.0%; Score 365; DB 4; Length 364;
Best Local Similarity 100.0%; Pred. No. 8e-39;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRSCGGVVTTRRRQCNPRPAGGRACVGADLQAEMLQCEKTLQLEFM 60

Db 59 WSSWGRSPCSRSCGGVVTTRRRQCNPRPAGGRACVGADLQAEMLQCEKTLQLEFM 118

```

RN      [1]
RP      SEQUENCE FROM N.A.
RA      Strausberg R.;
RL      Submitted (JUL-2002) to the EMBL/GenBank/DBDJ databases.
RL      EMBL; BC034739; AAH34739.1; -
DR      MGD; MGI:2429637; Adamts16;
DR      GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001590; Peptidase_M12B.
DR      InterPro; IPR002870; Peptidase_M12B_N.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      Pfam; PF01562; Pep_M12B_propep; 1.
DR      Pfam; PF01421; Reprolysin; 1.
DR      Pfam; PF00090; tsp_1; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00209; TSP1; 1.
DR      PROSITE; PS00215; ADAM_MPEPRO; 1.
DR      PROSITE; PS0092; TSP1; 1.
DR      Hypothetical protein.
DR      KW
SQ      SEQUENCE 900 AA; 100679 MW; 590BE2B0E73740FF CRC64;

Query Match          54.0%; Score 197; DB 11; Length 900;
Best Local Similarity 48.4%; Pred.No.9.1e-17;
Matches 31; Conservative 11; Mismatches 22; Indels 0; Gaps 0;

QY      1 WSSWGRSPSCRS CGGVVTRRQCNNRPFAFGGRACVAGDLQAEMCNTQACEKTQLQEFM 60
DB      587 WDSWSFWSPCSRCTCGGGISHRDLCNTNPRPSHGKFCQGSTRTIKLCNSQRCLDSVDFR 646
QY      61  SQQC 64
DB      647  AAQC 650

RESULT 4
Q8BKAI
ID      Q8BKAI      PRELIMINARY;      PET; 1092 AA.
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Hypothetical metalloprotease.
GS      E10314N14RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX      NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6J; TISSUE=Eye;
RC      MEDLINE=22354683; PubMed=12466851;
RA      The PANTOM Consortium.
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RL      60,770 full-length cDNAs."
RL      Nature 420:563-573 (2002).
DR      EMBL; AK053851; BAC35556.1; -
DR      MGD; MGI:2442620; E10314N14RIK.
DR      GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001590; Peptidase_M12B.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      Pfam; PF01421; Reprolysin_1.
DR      Pfam; PF00090; tsp_1; 5.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00209; TSP1; 5.
DR      PROSITE; PS00215; ADAM_MPEPRO; 1.
DR      PROSITE; PS0092; TSP1; 4.
DR      Hypothetical protein.
SQ      SEQUENCE 1092 AA; 120324 MW; B5F03956553536AE CRC64;

```

```

Query Match          51.2%; Score 187; DB 11; Length 1092;
Best Local Similarity 45.3%; Pred. No. 2.2e-15;
Matches 29; Conservative 13; Mismatches 22; Indels 0; Gaps 0;

Qy 1 WSSGPRSCSCGGVTRRQCNPPAFGGRACVGDLOAEMCNTOACEKTOLEFM 60
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 WSAWSKSCSCRTGGGVFQERHCSNPKPQYGGKYCPGSSRIYKLCNINPCNSLDFR 479
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

Qy 61 SQQC 64
Db 480 AQQC 483

RESULT 5
Q8MRL5 PRELIMINARY; PRT; 769 AA.
ID Q8MRL5
AC Q8MRL5
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE GH22104p.
GN CG4096.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Paablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hoston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banson J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferrier S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svitskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,

```



```

RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak P., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.,
RA "Annotation of Drosophila melanogaster genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003435; AAF46065.2; -.
DR HSP; F15167; IATL.
DR MEROPS; M12.231; -.
DR FlyBase; FBgn002791; CG4096.
DR GO; GO:0004222; Fmetalloendopeptidase activity; IEA.
DR GO; GO:0008270; Zinc ion binding; IEA.
DR GO; GO:0006508; Proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR006085; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp 1; 2.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 2.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSP1; 2.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1059 AA; 119195 MW; AD3C99B47618F3D7 CRC64;

Query Match 47.4%; Score 173; DB 5; Length 1059;
Best Local Similarity 45.3%; Pred. No. 1.3e-13;
Matches 29; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRSGGGVTRRRQCNPRPAFGGACVGDALQAEMLNTQACEKTLQLEPM 60
DB 631 WGDWSEWSECSRSGGGVTRRRQCNPRPAFGGACVGDALQAEMLNTQACEKTLQLEPM 60

QY 61 SQQC 64
DB 691 AQQC 694

RESULT 7
Q96SC3 PRELIMINARY; PRT; 2673 AA.
AC Q96SC3 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Fibulin-6 (Fragment).
GN FIBL-6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kostka G., Timpl R.;
RC TISSUE=Melanoma;
RA "Partial sequence of fibulin-6 with a c-terminal region related to
RT domain II and III of the fibulin family."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ306906; CAC37630.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.

```

```

DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000875; Cecropin.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR009017; GFP_Like.
DR InterPro; IPR007110; IG_Like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP1.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00047; IG; 17.
DR Pfam; PF00090; tsp 1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00179; EGF_CA; 7.
DR SMART; SM00408; IGC2; 17.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS00268; CECROPIN; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 8.
DR PROSITE; PS0835; IG_LIKE; 17.
DR PROSITE; PS50092; TSP1; 6.
KW EGF-like domain; Immunoglobulin domain.
FT NON_TER 1
SQ SEQUENCE 2673 AA; 291017 MW; BEAEC30B8340E272 CRC64;

Query Match 47.1%; Score 172; DB 4; Length 2673;
Best Local Similarity 55.8%; Pred. No. 4.5e-13;
Matches 29; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRSGGGVTRRRQCNPRPAFGGACVGDALQAEMLNTQAC 52
DB 1684 WSAWQPWGTCSESGKGTQTRALCNPPPAFGGSCDGAETQMVCNERNC 1735

RESULT 8
Q96RW7 PRELIMINARY; PRT; 5636 AA.
AC Q96RW7 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE Hemikentin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Trent J.;
RT "Human hemikentin gene."
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF156100; AAK68690.1; -.
DR GO; GO:0005727; C:extrachromosomal circular DNA; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004197; P:cysteine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000152; Asx_hydroxyl_S.
DR InterPro; IPR000875; Cecropin.
DR InterPro; IPR001434; DUF11.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR009017; GFP_Like.
DR InterPro; IPR007110; IG_Like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000169; SHprot_acsite.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00047; IG; 44.
DR Pfam; PF00090; tsp 1; 6.
DR SMART; SM00179; EGF_CA; 7.
DR SMART; SM00408; IGC2; 43.

```

```
DR SMART; SM00209; TSPL1; 6.
DR SMART; SM00327; VWA; 1.
DR TIGRFAMs; TIGR01451; B_ant_repeat; 9.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS00268; CECROPIN; 1.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 8.
DR PROSITE; PS00835; IG_LIKE; 44.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
DR PROSITE; PS0092; TSPL1; 6.
DR KW EGF-like domain; Immunoglobulin domain.
SQ SEQUENCE 5636 AA; 613660 MW; F00CB319CED7B52C CRC64;

Query Match 47.1%; Score 172; DB 4; Length 5636;
Best Local Similarity 55.8%; Pred. No. 9.3e-13;
Matches 29; Conservative 5; Mismatches 18; Indels 0; Gaps 0;

QY 1 WSSMGPRSCRSRGCGGVVTRRRQCNPRPAFGGRACVAGDLQAEVCNTOAC 52
DB 4647 WSAQWPWGTCSCGCGTQTRALCNPPPAFGGSGYCDGAETQVQVCNRC 4698

RESULT 9
Q8CG28 PRELIMINARY; PRT; 1070 AA.
AC Q8CG28;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Zinc metalloendopeptidase.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Brathwaite M.; Nagaraja R.; Abe K.;
RT "Genomic Sequence Analysis in the Mouse t-complex Region.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF528163; AA01380.1; -
DR GO; GO:0004222; F:metallopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B_N.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp_1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSPL1; 5.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSPL1_5.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
SQ SEQUENCE 1070 AA; 117203 MW; 6B896425EBA2B8D8 CRC64;

Query Match 46.6%; Score 170; DB 11; Length 1070;
Best Local Similarity 43.8%; Pred. No. 3.3e-13;
Matches 28; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 WSSMGPRSCRSRGCGGVVTRRRQCNPRPAFGGRACVAGDLQAEVCNTOACEKTOLEFM 60
DB 516 WGPWTPWGDCRSRGCGGVSSSRHRCDSPTTIGGKCYLGERRRHRCNTNDCPPGQDFR 575

QY 61 SQQC 64
DB 576 EMQC 579

RESULT 10
```

```
Q8SXB0 PRELIMINARY; PRT; 1688 AA.
ID Q8SXB0; Q9VF61;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GH16393p (CG14869 protein).
GN CG6107
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M.; Brokstein P.; Hong L.; Agbayani A.; Carlson J.;
RA Champe M.; Chavez C.; Dorsett V.; Dresnek D.; Farfan D.; Frise E.;
RA George R.; Gonzalez M.; Guarin H.; Kronmiller B.; Li P.; Liao G.;
RA Miranda A.; Mungall C.J.; Nunco J.; Pacleb J.; Paragas V.; Park S.;
RA Patel S.; Phouaneavong S.; Wan K.; Yu C.; Lewis S.E.; Rubin G.M.;
RA Celniker S.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D.; Celniker S.E.; Holt R.A.; Evans C.A.; Gocayne J.D.;
RA Amanatides P.G.; Scherer S.E.; Li P.W.; Hoskins R.A.; Galle R.F.;
RA George R.A.; Lewis S.E.; Richards S.; Ashburner M.; Henderson S.N.;
RA Sutton G.G.; Wortman J.R.; Yandell M.D.; Zhang Q.; Chen L.X.;
RA Brandon R.C.; Rogers Y.-H.C.; Blazej R.G.; Champe M.; Pfeiffer B.D.;
RA Wan K.H.; Doyle C.; Baxter E.G.; Helt G.; Nelson C.R.; Makos G.D.G.;
RA Abril J.F.; Agbayani A.; An H.-J.; Andrews-Pfannkoch C.; Baldwin D.;
RA Balow R.M.; Basu A.; Baxendale J.; Bayraktaroglu L.; Beasley E.M.;
RA Beeson K.Y.; Benos P.V.; Berman B.P.; Bhandari D.; Bolshakov S.;
RA Borkova D.; Botchan M.R.; Bouck J.; Brokstein P.; Brotter P.;
RA Burtis K.C.; Busam D.A.; Butler H.; Cadieu E.; Center A.; Chandra I.;
RA Cherry J.M.; Cawley S.; Dahlke C.; Davenport L.B.; Davies P.;
RA de Pablos B.; Delcher A.; Deng Z.; Mays A.D.; Dew I.; Dietz S.M.;
RA Dodson K.; Doup L.E.; Downes M.; Dugan-Rocha S.; Dunkov B.C.; Dunn P.;
RA Durbin K.J.; Evangelista C.C.; Ferraz C.; Ferrieri S.; Fleischmann W.;
RA Fessler C.; Gabriellian A.E.; Garg N.S.; Gelbart W.M.; Glasser K.;
RA Gilek A.; Gong F.; Gorrell J.H.; Gu Z.; Guan P.; Harris M.;
RA Harris N.L.; Harvey D.; Heiman T.J.; Hernandez J.R.; Houck J.;
RA Hostin D.; Houston K.A.; Howland T.J.; Wei M.-H.; Ibegwam C.;
RA Jalali M.; Kalush E.; Karpen G.H.; Ke Z.; Kennison J.A.; Ketchum K.A.;
RA Kimmel B.E.; Kodira C.D.; Kraft C.; Kravitz S.; Kulp D.; Lai Z.;
RA Lasko P.; Lei Y.; Levitsky A.A.; Li J.; Li Z.; Liang Y.; Lin X.;
RA Liu X.; Mattei B.; McIntosh T.C.; McLeod M.P.; McPherson D.;
RA Merkulov G.; Milshina N.V.; Mobarry C.; Morris J.; Moshrefi A.;
RA Mount S.M.; Moy M.; Murphy B.; Murphy L.; Muzny D.M.; Nelson D.L.;
RA Nelson D.R.; Nelson K.A.; Nixon K.; Nusskern D.R.; Pacleb J.M.;
RA Palazzolo M.; Pittman G.S.; Pan S.; Pollard J.; Puri V.; Reese M.G.;
RA Reinert K.; Remington K.; Saunders R.D.C.; Scheeler F.; Shen H.;
RA Shue B.C.; Siden-Kiamos I.; Simpson M.; Skupski M.P.; Smith T.;
RA Spier E.; Spradling A.C.; Stapleton M.; Strong R.; Sun E.;
RA Svirskas R.; Tector C.; Turner R.; Venter E.; Wang A.H.; Wang X.;
RA Wang Z.-Y.; Wassarman D.A.; Weinstock G.M.; Weissbach J.;
RA Williams S.M.; Woodage T.; Worley K.C.; Wu D.; Yang S.; Yao Q.A.;
RA Ye J.; Yeh R.-F.; Zaveri J.S.; Zhan M.; Zhang G.; Zhao Q.; Zheng L.;
RA Zheng X.H.; Zhong F.N.; Zhong W.; Zhou X.; Zhu S.; Zhu X.; Smith H.O.;
RA Gibbs R.A.; Myers E.W.; Rubin G.M.; Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:12185-2195(2000).
RN [3]
RP SEQUENCE FROM N.A.
RA Celniker S.E.; Adams M.D.; Kronmiller B.; Wan K.H.; Holt R.A.;
RA Evans C.A.; Gocayne J.D.; Amanatides P.G.; Brandon R.C.; Rogers Y.;
RA Banzon J.; An H.; Baldwin D.; Banzon J.; Beeson K.Y.; Busam D.A.;
RA Carlson J.W.; Center A.; Champe M.; Davenport L.B.; Dietz S.M.;
RA Dodson K.; Dorsett V.; Doup L.E.; Doyle C.; Dresnek D.; Farfan D.;
RA Ferreira S.; Frise E.; Galle R.F.; Garg N.S.; George R.A.;
```

RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RA "Sequencing of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bergman C., Carlson J.W., Celisner S.E.,
RA Clamp J., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [6]
RP SEQUENCE FROM N.A.
RA FlyBase,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY094716; AAW1069.1; -
DR EMBL; AE003709; AAF55199.2; -
DR FlyBase; FBgn0038340; CG6107.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0008270; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B_N.
DR InterPro; IPR006025; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp 1; 4.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR NON_TER 1009 1009
SQ SEQUENCE 1009 AA; 113352 MW; 7B63218CFE0FDB1 CRC64;
Query Match 45.8%; Score 167; DB 11; Length 1009;
Best Local Similarity 42.2%; Pred. No. 7.6e-13;
Matches 27; Conservative 5; Mismatches 32; Indels 0; Gaps 0;
QY 1 WSWGPRSPCSRSRGGGVVTRRRQCNPRPAFGGRACVGDLOAEMCNTOACEKTOLEFM 60
DB 549 WGRWSPWHSCHSRTCGAGAAQLNNPFRFGKCYCTGERKRYLNCNVHPCRSDTPFR 608
QY 61 SQQC 64
DB 609 QMOC 512
RESULT 12
Q811B3 PRELIMINARY; PRT; 1600 AA.
ID Q811B3
AC Q811B3
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metalloprotease disintegrin 12 protein.
GN ADAMTS-12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Cal S., Lopez-Otin C.;
RA "Mouse ADAMTS-12";
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AJ537452; CAD60967.1; -
DR PIR; PT0546; PT0698.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B_N.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR000885; TSP1.

RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RA "Sequencing of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [4]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bergman C., Carlson J.W., Celisner S.E.,
RA Clamp J., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of Drosophila melanogaster genome";
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [5]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celisner S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL [6]
RP SEQUENCE FROM N.A.
RA FlyBase,
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY094716; AAW1069.1; -
DR EMBL; AE003709; AAF55199.2; -
DR FlyBase; FBgn0038340; CG6107.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0008270; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase_M12B_N.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp 1; 2.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR SEQUENCE 1688 AA; 189867 MW; 48FB8DD4DE0CA4D2 CRC64;
Query Match 46.0%; Score 168; DB 5; Length 1688;
Best Local Similarity 43.8%; Pred. No. 9.4e-13;
Matches 28; Conservative 8; Mismatches 28; Indels 0; Gaps 0;
QY 1 WSWGPRSPCSRSRGGGVVTRRRQCNPRPAFGGRACVGDLOAEMCNTOACEKTOLEFM 60
DB 754 WGPWTPTPCSLTCGGGVQESRRECQFVPENGKCYCTGSRKRYRSCNTHQCPFGSMDPR 813
QY 61 SQQC 64
DB 814 EQQC 817
RESULT 11
Q8BKX1 PRELIMINARY; PRT; 1009 AA.
ID Q8BKX1
AC Q8BKX1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ADAMTS-12 precursor (Fragment).
GN A1605170.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; TSP 1; 7.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSP1; 8.
 DR PROSITE; PS50215; ADAM_MPRO; 1.
 DR PROSITE; PS50092; TSP1; 6.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Integrin; Protease.
 SQ SEQUENCE 1600 AA; 177791 MW; 083089D35E38C5D CRC64;
 Query Match 45.8%; Score 167; DB 11; Length 1600;
 Best Local Similarity 42.2%; Pred. No. 1.2e-12;
 Matches 27; Conservative 5; Mismatches 32; Indels 0; Gaps 0;
 QY 1 WSWGPRSPCSRSCGGVWTRRQCNPNRPAFGRCVADLQAEQNTQACEKTLQEFM 60
 DB 549 WGRWSPWHSCTCGAGQAERLCNPNPKFGKYCTGKERKRYLCNVHPCRSDTPTFR 608
 QY 61 SQQC 64
 DB 609 QMQC 612
 RESULT 13
 ID Q9GL54 PRELIMINARY; PRT; 269 AA.
 AC Q9GL54;
 DT 01-MAR-2001 (TReMBLrel. 16, Created)
 DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Aggrecanase-2 (Fragment).
 GN ADAMTS-11.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA "Molecular cloning of lapine aggrecanase-2 (ADAMTS-11) from articular chondrocytes."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF317415; AAG33062.1; -.
 DR HSSP; Q9PW35; 1BUD.
 DR MEROPS; M12.225; -.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000895; TSP 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00209; TSP1; 1.
 DR PROSITE; PS50215; ADAM_MPRO; 1.
 DR PROSITE; PS50092; TSP1; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 FT NON_TER 1 1
 FT NON_TER 269 269
 SQ SEQUENCE 269 AA; 29193 MW; 97A1CA80B33452FA CRC64;
 Query Match 44.1%; Score 161; DB 6; Length 269;
 Best Local Similarity 45.3%; Pred. No. 1.2e-12;
 Matches 29; Conservative 2; Mismatches 33; Indels 0; Gaps 0;
 QY 1 WSWGPRSPCSRSCGGVWTRRQCNPNRPAFGRCVADLQAEQNTQACEKTLQEFM 60
 DB 193 WSWGPRSPCSRSCGGVWTRRQCNPNRPAFGRCVADLQAEQNTQACEKTLQEFM 322
 QY 61 SQQC 64

Db 253 HEQC 256
 RESULT 14
 ID Q8BGP4 PRELIMINARY; PRT; 623 AA.
 AC Q8BGP4;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE A disintegrin-like and metalloprotease (Fragment).
 GN ADAMTS5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath.; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Retina;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."
 RT Nature 420:563-573(2002).
 RL Nature 420:563-573(2002).
 DR EMBL; AK044746; BAC32062.1; -.
 DR EMBL; AK082475; BAC38503.1; -.
 DR MGD; MGI:1346321; Adamts5.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0008318; F:protein prenyltransferase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0018346; P:protein amino acid prenylation; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR006586; ADAM cysteine
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR002088; PPTA.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR000895; TSP 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR PRINTS; PR01705; TSP1REPEAT.
 DR SMART; SM00608; ACR; 1.
 DR SMART; SM00209; TSP1; 2.
 DR PROSITE; PS50215; ADAM_MPRO; 1.
 DR PROSITE; PS00904; PPTA; 1.
 DR PROSITE; PS50092; TSP1; 2.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 FT NON_TER 1 1
 FT NON_TER 623 AA; 68681 MW; 80CE9BCD0AC37502 CRC64;
 SQ SEQUENCE 623 AA; 68681 MW; 80CE9BCD0AC37502 CRC64;
 Query Match 44.1%; Score 161; DB 11; Length 623;
 Best Local Similarity 45.3%; Pred. No. 2.8e-12;
 Matches 29; Conservative 2; Mismatches 33; Indels 0; Gaps 0;
 QY 1 WSWGPRSPCSRSCGGVWTRRQCNPNRPAFGRCVADLQAEQNTQACEKTLQEFM 60
 DB 263 WSWGPRSPCSRSCGGVWTRRQCNPNRPAFGRCVADLQAEQNTQACEKTLQEFM 322
 QY 61 SQQC 64
 DB 323 HEQC 326
 RESULT 15
 ID Q9U631 PRELIMINARY; PRT; 1081 AA.
 AC Q9U631;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
 DE Semaphorin 5C protein.

GN SEMA-SC OR CG5661.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R;
RA Khare N., Daroch S., Chiquet-Ehrismann R., Baumgartner S.;
RT "Expression patterns of two new members of the semaphorin family in
RL Drosophila suggest early functions during embryogenesis.";
DR Mech. Dev. 0:0-0(1999);
DR EMBL; AF198084; AAF04860.1; -;
DR Flybase; FBgn0028679; SEMA-SC.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; P:catalytic activity; IEA.
DR GO; GO:0008199; P:feric iron binding; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0006725; P:aromatic compound metabolism; IEA.
DR GO; GO:0007275; P:development; IEA.
DR InterPro; IPR001893; Decarboxylase2.
DR InterPro; IPR000627; Dioxigenase.
DR InterPro; IPR003659; Flexin-like.
DR InterPro; IPR002165; Flexin-repeat.
DR InterPro; IPR001627; SEMA.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; SEMA; 1.
DR Pfam; PF00090; tsp_1; 6.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; SEMA; 1.
DR SMART; SM00203; TSP1; 6.
DR PROSITE; PS00879; ODR_DC_2; 1.
DR PROSITE; PS00092; TSP1; 6.
SQ SEQUENCE 1081 AA; 120435 MW; 561071C831C431D3 CRC64;

Query Match 44.1%; Score 161; DB 5; Length 1081;
Best Local Similarity 48.2%; Pred. No. 4.8e-12;
Matches 27; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSGCGGVVTRRQCNNRPAFGGRACVGADLQAEKNTQACEKQ 56
Db 662 WGPWGEWSECSAQCGGGRMRRECDNPAPNGMECPGRLDYECNMOSQCEVR 717

Search completed: March 13, 2004, 07:43:19
Job time : 5.33333 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 9.36667 Seconds
(without alignments)
1930.574 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473

Perfect score: 365
Sequence: 1 WSSWGPRAPCSRSGGGVVT.....EMCNTQACEKTLQLEFMSQQC 64

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*

- 1: Geneseq1980s.*
- 2: Geneseq1990s.*
- 3: Geneseq2000s.*
- 4: Geneseq2001s.*
- 5: Geneseq2002s.*
- 6: Geneseq2003as.*
- 7: Geneseq2003bs.*
- 8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	365	100.0	364	5	Aau97642 Human Hsa
2	365	100.0	364	5	Abp43989 Procollag
3	365	100.0	365	7	Add94035 Human Hsa
4	365	100.0	933	5	Aau79217 Human ADA
5	365	100.0	1297	6	Aao16618 Human von
6	365	100.0	1323	6	Aao16620 Human von
7	365	100.0	1353	5	Aae24449 Human von
8	365	100.0	1353	6	Aao16617 Human von
9	365	100.0	1378	6	Aao16619 Human von
10	365	100.0	1416	5	Abd04153 Human ADA
11	365	100.0	1427	5	Aae24450 Human von
12	365	100.0	1427	7	Add94038 Human agg
13	365	100.0	1445	5	Abb98125 Human PMX
14	206	56.4	700	6	Ada50467 Human pro
15	206	56.4	1048	4	Abb5695 Larynx ca
16	206	56.4	1156	5	Abg30864 Human met
17	206	56.4	1186	5	Aau72895 Human met
18	206	56.4	1189	5	Aau74750 Human pro
19	206	56.4	1210	5	Aau85415 Human pro
20	206	56.4	1213	5	Aau77128 Human pro
21	206	56.4	1216	5	Aau77127 Human pro
22	206	56.4	1219	5	Aau77126 Human pro
23	206	56.4	1222	5	Aau77125 Human pro
24	206	56.4	1223	5	Abg30863 Human met
25	206	56.4	1223	5	Aau79747 Human ADA

Abp71574 Human agg
Abp42736 Human ADA
Aau77130 Human pro
Abp71573 Human agg
Aau77129 Human pro
Aau77132 Human pro
Aau77131 Human pro
Abp71575 Human agg
Abp76895 Human zln
Abg31503 Human ADA
Ada50469 Human pro
Abg76897 Human ADA
Aau79496 Human par
Abp40093 Human ADA
Aau79500 Human ADA
Aau72891 Human met
Aao15254 Human MDT
Aau79497 Human ADA
Aau80153 Human ADA
Abg72431 Novel hum

ALIGNMENTS

RESULT 1

AAU97642
ID AAU97642 standard; protein; 364 AA.

XX AC AAU97642;

XX DT 27-AUG-2002 (first entry)

XX DE Human Hsa011374 protein.

XX KW Aggrecanase; osteopathic; antiarthritic; antiinflammatory; human;
XX KW cartilage; osteoarthritis; inflammatory disease; enzyme; Hsa011374.

XX OS Homo sapiens.

XX PN WO200233093-A2.

XX PD 25-APR-2002.

XX PF 17-OCT-2001; 2001WO-US032458.

XX PR 18-OCT-2000; 2000US-0241469P.

XX (GEM) GENETICS INST INC.

XX PI Racie LA, Twine NC, Agostino MJ, Wolfman NM, Morris EA;

XX DR WPI; 2002-454602/48.

XX DR N-PSDB; ABK52580.

XX PT Novel purified aggrecanase polypeptide useful for developing inhibitors
XX PT and antibodies to the aggrecanase polypeptide, which are useful for
XX PT treating aggrecanase-associated condition such as osteoarthritis.

XX PS Disclosure; Page 37-38; 41pp; English.

XX CC This invention relates to the cDNA and protein sequences of a novel human
XX CC aggrecanase polypeptide. The protein of the invention may be used to
XX CC inhibit the proteolytic activity of aggrecanase, or to inhibit the
XX CC aggrecanase-mediated cleavage of aggrecan in cartilage. The protein of
XX CC the invention is useful for developing inhibitors of aggrecanase protein.
XX CC The cDNA sequence encoding the aggrecanase protein of the invention is
XX CC useful for designing probes for obtaining DNA sequences encoding other
XX CC aggrecanase molecules. The cDNA sequence is also useful for detecting
XX CC mRNA encoding aggrecanase in a given cell population, and thus for
XX CC detecting or diagnosing genetic disorders involving the aggrecanase, or
XX CC disorders involving cellular, organ or tissue disorders in which
XX CC aggrecanase is irregularly transcribed or expressed. The DNA sequences

CC may also be useful for preparing vectors for gene therapy applications.
 CC An inhibitor of the protein is useful in treating conditions
 CC characterised by degradation of articular cartilage, by blocking the
 CC enzyme's proteolytic activity. An aggrecanase protein inhibitor and a
 CC method for inhibition of its activity are useful for treating various
 CC aggrecanase-associated conditions including osteoarthritis and other
 CC inflammatory diseases. The present sequence represents the human
 CC Hsa011374 protein of the invention
 XX SQ Sequence 364 AA;
 Query Match 100.0%; Score 365; DB 5; Length 364;
 Best Local Similarity 100.0%; Pred. No. 3.3e-31;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPAGGRACVADLQAEVCNTOACEKTQLBFM 60
 Db 59 WSSWGRSPCSRSCGGVVTTRRQCNPRPAGGRACVADLQAEVCNTOACEKTQLBFM 118
 Qy 61 SQQC 64
 Db 119 SQQC 122
 RESULT 2
 ID ABP43989 standard; protein; 364 AA.
 AC ABP43989;
 DT 26-FEB-2003 (first entry)
 DE Procollagen I N-proteinase.
 XX Neuroprotective; immunomodulator; cancer; chromosome 9p34; cytostatic;
 KW anti-inflammatory; Gene therapy; nutritional supplement; wound; burn;
 KW ulcer; Alzheimer's disease; Huntington's disease;
 KW amyotrophic lateral sclerosis; autoimmune disorder; inflammation;
 KW vulnary.
 OS Bos taurus.
 XX WO200231111-A2.
 XX 18-APR-2002.
 XX 11-OCT-2001; 2001WO-US027760.
 XX 12-OCT-2000; 2000US-00687527.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 XX WPI; 2002-426278/45.
 DR N-PSDB; ABQ61233.
 XX New polypeptides and their encoded proteins, useful as nutritional
 PT sources or supplements, or in gene therapy, particularly for treating
 PT wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or
 PT inflammation.
 XX Claim 20; SEQ ID # 892; 357pp + Sequence Listing; English.
 PS The invention relates to 446 newly isolated polynucleotide sequences. The
 CC activity of polynucleotides of the invention may be described as,
 CC vulnary, neuroprotective, immunomodulator, cytostatic and anti-
 CC inflammatory. Compositions comprising nucleic acids of the invention are
 CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or

CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records ABP4344-
 CC ABP43989 represent polypeptides encoded by polynucleotides of the
 CC invention. NOTE: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 364 AA;
 Query Match 100.0%; Score 365; DB 5; Length 364;
 Best Local Similarity 100.0%; Pred. No. 3.3e-31;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPAGGRACVADLQAEVCNTOACEKTQLBFM 60
 Db 59 WSSWGRSPCSRSCGGVVTTRRQCNPRPAGGRACVADLQAEVCNTOACEKTQLBFM 118
 Qy 61 SQQC 64
 Db 119 SQQC 122
 RESULT 3
 ID ADD94035 standard; protein; 365 AA.
 AC ADD94035;
 DT 29-JAN-2004 (first entry)
 DE Human Hsa011374 protein 1 related to aggrecanase.
 XX aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
 KW cartilage degradation; osteoarthritis; inflammatory joint disease;
 KW antarthritic; osteopathic; antiinflammatory;
 KW aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
 KW human; Hsa011374.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH Misc-difference 365 /label= OTHER
 FT /note= Unknown amino acid; Encoded by TGA"
 FT
 XX US2003105313-A1.
 XX 05-JUN-2003.
 XX 25-JAN-2002; 2002US-00057487.
 XX 16-OCT-2001; 2001US-00978979.
 XX (AMHP) AMERICAN HOME PROD CORP.
 XX Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
 PI WPI; 2003-801251/75.
 DR N-PSDB; ADD94034.
 XX New isolated DNA molecule encoding an aggrecanase polypeptide for
 PT producing a purified human aggrecanase protein which can be used to
 PT develop inhibitors of aggrecanase.
 XX Disclosure; SEQ ID NO 5; 24pp; English.
 XX This invention relates to a novel human aggrecanase protein and the DNA
 CC sequence which encodes it. Aggrecan is a major extracellular component of
 CC articular cartilage. It is a proteoglycan responsible for providing
 CC cartilage with its mechanical properties of compressability and
 CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
 CC cleavage of aggrecan thereby having a role in cartilage degradation
 CC associated with osteoarthritis and inflammatory joint disease. Compounds

CC which inhibit the activity of the protein of the invention may have
 CC antiarthritic, osteopathic or antiinflammatory activity. The invention
 CC may be used to produce a purified human aggrecanase protein. The protein
 CC (or fragment) may be used to develop inhibitors of aggrecanase using
 CC three dimensional structural analysis or computer aided drug design. A
 CC peptide which binds to aggrecanase is used to inhibit the proteolytic
 CC degradation of aggrecan. The invention may be useful for the development
 CC of therapeutics for the treatment of aggrecanase-associated disorders,
 CC such as, osteoarthritis and other inflammatory conditions. The present
 CC sequence is the amino acid sequence of the Hs011374 protein 1 which is
 CC related to the human aggrecanase protein of the invention.

XX
 SQ Sequence 365 AA;

Query Match 100.0%; Score 365; DB 7; Length 365;
 Best Local Similarity 100.0%; Pred. NO. 3.3e-31; Indels 0; Gaps 0;
 Matches 64; Conservative 0; Mismatches 0;

Qy 1 WSSWGPRSPCSRSCGGVTRRQCNPRPAFGGRACVGDADLAEMCNTOACEKTQLEFM 60
 Db 59 WSSWGPRSPCSRSCGGVTRRQCNPRPAFGGRACVGDADLAEMCNTOACEKTQLEFM 118
 Qy 61 SQQC 64
 Db 119 SQQC 122

RESULT 4
 AAU79217
 ID AAU79217 standard; protein; 933 AA.

XX AC AAU79217;

DT 15-JUL-2002 (first entry)

DE Human ADAM-TS-like protein.

XX Human; ADAM-TS-like protein; cardiovascular disorder; angina;

KW vascular system; congestive heart failure; myocardial infarction;

KW ischaemic heart disease; arrhythmia; hypertensive vascular disease;

KW secondary arterial hypertension; peripheral vascular disease; embolism;

KW chronic peripheral arterial occlusive disease; acute arterial thrombosis;

KW inflammatory vascular disorder; chronic obstructive pulmonary disease;

KW liver disorder.

OS Homo sapiens.

XX WO200226999-A2.

PN 04-APR-2002.

XX 26-SEP-2001; 2001WO-EP011124.

XX 28-SEP-2000; 2000US-0235881P.

XX 25-JUL-2001; 2001US-0307393P.

XX (FARB) BAYER AG.

XX Xiao Y;

PI WPI; 2002-383274/41.

XX N-PSDB; ABX49356.

XX New purified human ADAM-TS-like protein, useful for identifying

PT modulators of protein activity for treating cardiovascular or liver

PT disorder or chronic obstructive pulmonary disease.

XX Claim 25; Fig 2; 106pp; English.

XX The invention relates to a human ADAM-TS-like protein and the

CC polynucleotide encoding it. The protein of the invention is useful for

CC treating cardiovascular disorders including diseases of the heart and

CC vascular system, such as congestive heart failure, myocardial infarction,

CC ischaemic heart diseases (e.g., stable angina, unstable angina), atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases (e.g., all
 CC kinds of secondary arterial hypertension), and peripheral vascular
 CC diseases (e.g., chronic peripheral arterial occlusive disease, acute
 CC arterial thrombosis and embolism, inflammatory vascular disorders, etc),
 CC liver disorders and chronic obstructive pulmonary disease. The sequences
 CC are useful in diagnostic assays for detecting diseases and abnormalities
 CC or susceptibility to diseases and abnormalities related to the presence
 CC of mutations in the nucleic acid sequences which encode the protein. The
 CC sequences are also useful for modulating ADAM-TS-like protein activity in
 CC a disease condition. This sequence represents the human ADAM-TS-like
 CC protein

XX Sequence 933 AA;

Query Match 100.0%; Score 365; DB 5; Length 933;
 Best Local Similarity 100.0%; Pred. NO. 8.3e-31;
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSRSCGGVTRRQCNPRPAFGGRACVGDADLAEMCNTOACEKTQLEFM 60
 Db 390 WSSWGPRSPCSRSCGGVTRRQCNPRPAFGGRACVGDADLAEMCNTOACEKTQLEFM 449
 Qy 61 SQQC 64
 Db 450 SQQC 453

RESULT 5

AAO16618

ID AAO16618 standard; protein; 1297 AA.

XX AC AAO16618;

XX 15-MAY-2003 (first entry)

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #5.

XX Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

KW platelet thrombosis; stenosis.

XX Homo sapiens.

XX WO200288366-A1.

XX 07-NOV-2002.

XX 25-APR-2002; 2002WO-JP004141.

XX 25-APR-2001; 2001JP-00128342.

XX 27-JUL-2001; 2001JP-00227510.

XX 28-SEP-2001; 2001JP-00302977.

XX 25-JAN-2002; 2002JP-00017596.

XX (KAGA) CHERO-SERO-THERAPEUTIC RES INST.

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX WPI; 2003-120479/11.

XX N-PSDB; ABT32584.

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
 PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
 PT developing drugs for e.g. myocardial infarction and cerebral infarction.

XX Claim 4; Page 92-101; 144pp; Japanese.

XX The invention comprises the amino acid and coding sequence of a von

CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of

CC the invention are useful in the diagnosis and treatment of

CC thrombocytopenic purpura, and in developing drugs for myocardial

CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein
XX
SQ Sequence 1297 AA;

Query Match 100.0%; Score 365; DB 6; Length 1297;
Best Local Similarity 100.0%; Pred. No. 1.2e-30; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;

Qy 1 WSSWGPRSPCSRSCGGVVTTRRQCNPNRPAFGGRACVADLOAEMCNTOACEKTQLEFM 60
Db 313 WSSWGPRSPCSRSCGGVVTTRRQCNPNRPAFGGRACVADLOAEMCNTOACEKTQLEFM 372
Qy 61 SQQC 64
Db 373 SQQC 376

RESULT 6
AAOI16620
ID AAOI16620 standard; protein; 1323 AA.
AC AAOI16620;
XX
DT 15-MAY-2003 (first entry)
DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #7.
XX
KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
KW myocardial infarction; cerebral infarction; arteriosclerosis;
KW platelet thrombosis; stenosis.

OS Homo sapiens.
XX
PN WO200288366-A1.
XX
PD 07-NOV-2002.

XX 25-APR-2002; 2002WO-JP004141.
XX 25-APR-2001; 2001JP-00128342.
PR 27-JUL-2001; 2001JP-00227510.
PR 28-SEP-2001; 2001JP-00302977.
PR 25-JAN-2002; 2002JP-00017596.

XX (KAGA) CEMO-SERO-THERAPEUTIC RES INST.
XX
PI Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
XX
DR WPI; 2003-120479/11.
DR N-PSDB; ABT32586.

XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
PT developing drugs for e.g. myocardial infarction and cerebral infarction.
XX
PS Claim 4; Page 112-121; 144pp; Japanese.

XX The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein

XX
SQ Sequence 1323 AA;
Query Match 100.0%; Score 365; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 1.2e-30; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;

Qy 1 WSSWGPRSPCSRSCGGVVTTRRQCNPNRPAFGGRACVADLOAEMCNTOACEKTQLEFM 60
Db 338 WSSWGPRSPCSRSCGGVVTTRRQCNPNRPAFGGRACVADLOAEMCNTOACEKTQLEFM 397
Qy 61 SQQC 64
Db 398 SQQC 401

RESULT 7
AAE24449
ID AAE24449 standard; protein; 1353 AA.
XX
AC AAE24449;
XX
DT 04-OCT-2002 (first entry)
DE Human von Willebrand factor-cleaving protease fragment #2.

XX
KW Human; von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;
KW transgenic animal; immunisation; thromboembolic disease; preeclampsia;
KW thrombotic thrombocytopenic purpura; TTP; Henoch-Schonlein purpura;
KW thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;
KW transgenic; anticoagulant.

XX Homo sapiens.
OS
XX WO200242441-A2.
XX
PD 30-MAY-2002.

XX 20-NOV-2001; 2001WO-EP013391.
XX
PR 22-NOV-2000; 2000US-00721254.
PR 12-APR-2001; 2001US-00833328.

XX (BAXT) BAXTER AG.
XX
PI Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
PI Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
PI Zimmermann K, Voelkel D;
XX
DR WPI; 2002-479950/51.

XX Novel isolated or substantially purified von Willebrand factor-cleaving
PT protease, useful for producing preparation for therapy of thrombosis and
PT thromboembolic disease such as thrombotic thrombocytopenic purpura.
XX
PS Claim 1; Page 64-68; 93pp; English.

XX The invention relates to an isolated or substantially pure von Willebrand
CC factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
CC purifying vWF which involves providing vWF-cp as a ligand, contacting a
CC solution comprising vWF with the polypeptide ligand under conditions
CC where vWF is bound to the ligand and recovering from the ligand purified
CC vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
CC which involves immunising an animal with vWF-cp and isolating the anti-
CC vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
CC producing a preparation of prophylaxis and therapy of thrombosis and
CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
CC Henoch-Schonlein purpura, preeclampsia, neonatal thrombocytopenia or
CC haemolytic-uraemic syndrome. vWF-cp can also be used for processing
CC plasmatric or recombinantly produced vWF. The invention is useful for
CC constriction expression systems and generating transgenic animals which
CC express the polypeptide in vivo. The present sequence is human vWF-cp
CC fragment

XX
SQ Sequence 1353 AA;
Query Match 100.0%; Score 365; DB 5; Length 1353;
Best Local Similarity 100.0%; Pred. No. 1.2e-30; Indels 0; Gaps 0;
Matches 64; Conservative 0; Mismatches 0;

...

DE	XX	Human ADAMTS-M polypeptide.	
XX	XX	Osteoarthritis; rheumatoid arthritis; inflammatory bowel disease;	
KW	KW	Crohn's disease; asthma; Alzheimer's disease; organ transplant rejection;	
KW	KW	cachexia; allergy; cancer; leukaemia; lymphoma; osteoporosis;	
KW	KW	atherosclerosis; congestive heart failure; myocardial infarction; stroke;	
KW	KW	neurodegenerative disease; autoimmune disorder; Huntington's;	
KW	KW	Parkinson's; migraine; pain; depression; multiple sclerosis; burn;	
KW	KW	infertility; diabetic shock; gene therapy; ADAMTS-M;	
KW	KW	A Disintegrin And Metalloprotease; thrombospondin domain.	
XX	XX	Homo sapiens.	
XX	XX		
XX	XX	Key	Location/Qualifiers
XX	XX	Peptide	1..97
XX	XX		/label= Prodomain
XX	XX		/note= "The mature form of the ADAMTS-M protein is
XX	XX		processed by furin cleavage of the prodomain"
XX	XX	Cleavage-site	94..97
XX	XX		/label= Furin_cleavage_site
XX	XX	Protein	98..1416
XX	XX		/label= Mature ADAMTS-M protein
XX	XX		/note= "The mature form of the ADAMTS-M protein is
XX	XX		processed by furin cleavage of the prodomain"
XX	XX	Domain	98..311
XX	XX		/label= Metalloprotease_domain
XX	XX	Domain	247..272
XX	XX		/label= Zinc-binding_motif
XX	XX	Domain	324..394
XX	XX		/label= Disintegrin_domain
XX	XX	Domain	410..473
XX	XX		/label= Thrombospondin_submotif
XX	XX	Domain	419..424
XX	XX		/label= Heparin-binding_domain
XX	XX	Domain	1099..1156
XX	XX		/label= Thrombospondin_submotif
XX	XX	EP1152055-A1.	
XX	XX	07-NOV-2001.	
XX	XX	24-APR-2001; 2001EP-00303706.	
XX	XX	27-APR-2000; 2000US-0200040P.	
XX	XX	(PFIZ) PFIZER PROD INC.	
XX	XX	Buckbinder L, Mitchell PG, Wachtmann TS, Walsh RT;	
XX	XX	WPI; 2002-084275/12.	
XX	XX	N-PSDB; ABA02549.	
XX	XX	New polynucleotide, useful in gene therapy, particularly for treating or	
XX	XX	preventing e.g. arthritis, Crohn's disease, Alzheimer's disease and organ	
XX	XX	transplant toxicity and rejection, comprises ADAMTS polynucleotide and	
XX	XX	encoded polypeptide.	
XX	XX	Claim 4; Fig 2; 31pp; English.	
XX	XX	The present sequence represents a ADAMTS protein, designated ADAMTS-M,	
XX	XX	that exhibits the characteristics of the ADAM (A Disintegrin And	
XX	XX	Metalloprotease) family of metalloproteases, and contains a	
XX	XX	thrombospondin domain (TS). The protein is encoded by the cDNA given in	
XX	XX	ABA02549. The specification describes a newly isolated polynucleotide,	
XX	XX	comprising a nucleotide sequence encoding an ADAMTS-M polypeptide as	
XX	XX	given in the specification, or a metalloproteinase, disintegrin domain,	
XX	XX	prodomain or its thrombospondin submotif. The polynucleotide, polypeptide	
XX	XX	and agent are useful for manufacturing a medicament for treating a	
XX	XX	subject in need of altering activity or expression of ADAMTS-M. The	
XX	XX	polynucleotide, ADAMTS-M polypeptide and agent are useful for	
XX	XX	manufacturing a medicament for treating arthritis (osteoarthritis and	
XX	XX	rheumatoid arthritis), inflammatory bowel disease, Crohn's disease,	
XX	XX	asthma, Alzheimer's disease, organ transplant toxicity and rejection,	
CC	CC	cachexia, allergy, cancer (e.g. solid tumour cancer including colon,	
CC	CC	breast, lung, prostate, brain or haematopoietic malignancies including	
CC	CC	leukaemia and lymphoma), osteoporosis, atherosclerosis, aortic aneurysm,	
CC	CC	congestive heart failure, myocardial infarction, stroke, head trauma,	
CC	CC	spinal cord injury, neurodegenerative disease, autoimmune disorders,	
CC	CC	Huntington's disease, Parkinson's disease, migraine, pain, depression,	
CC	CC	multiple sclerosis, abnormal wound healing, burns, infertility or	
CC	CC	diabetic shock. The polynucleotide and polypeptide are also useful for	
CC	CC	diagnosing the diseases above. The polynucleotide is particularly useful	
CC	CC	in gene therapy for treating the diseases cited above	
XX	XX	Sequence 1416 AA;	
SQ	SQ		
		Query Match	100.0%; Score 365; DB 5; Length 1416;
		Best Local Similarity	100.0%; Pred. No. 1.3e-30;
		Matches	64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	WSSWGPRSPCSRS CGGVVTRRRQCNPRPAGGRACVGDADLAQMCMNTQACEKTQLEFM	60
Db	410	WSSWGPRSPCSRS CGGVVTRRRQCNPRPAGGRACVGDADLAQMCMNTQACEKTQLEFM	469
Qy	61	SQQC 64	
Db	470	SQQC 473	
		RESULT 11	
		AAE24450	
ID	AAE24450	standard; protein; 1427 AA.	
XX	XX	AAE24450;	
XX	XX	04-OCT-2002 (first entry)	
XX	XX	Human Von Willebrand factor-cleaving protease (vWF-cp).	
DE	XX	Human; Von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;	
KW	KW	transgenic animal; immunisation; thromboembolic disease; preclampsia;	
KW	KW	thrombotic thrombocytopenic purpura; TTP; Henoch-Schönlein purpura;	
KW	KW	thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;	
KW	KW	transgenic; anticoagulant; chromosome 9.	
XX	XX	Homo sapiens.	
XX	XX		
XX	XX	Key	Location/Qualifiers
XX	XX	Peptide	1..51
XX	XX		/label= Signal_peptide
XX	XX	Protein	52..1427
XX	XX		/note= "Mature human vWF-cp protein"
XX	XX	Cleavage-site	69..75
XX	XX		/note= "Furin cleavage site"
XX	XX	Region	224..228
XX	XX		/note= "Catalytical side"
XX	XX	Region	249
XX	XX		/note= "Met turn"
XX	XX	Domain	301..377
XX	XX		/note= "Disintegrin like motif"
XX	XX	Domain	387..439
XX	XX		/note= "Thrombospondin type I motif"
XX	XX	Region	441..553
XX	XX		/note= "Cysteine rich region"
XX	XX	Region	554..687
XX	XX		/note= "Spacer"
XX	XX	Domain	688..743
XX	XX		/note= "Thrombospondin type I motif"
XX	XX	Domain	744..805
XX	XX		/note= "Thrombospondin type I motif"
XX	XX	Domain	897..952
XX	XX		/note= "Thrombospondin type I motif"
XX	XX	Domain	953..1013
XX	XX		/note= "Thrombospondin type I motif"
XX	XX	Domain	1016..1073
XX	XX		/note= "Thrombospondin type I motif"

FT	Domain	1075..1131	
FT	/note= "Thromspondin type I motif"		
XX	PN	WO200242441-A2.	
XX	PD	30-MAY-2002.	
XX	PF	20-NOV-2001; 2001WO-EP013391.	
XX	PR	22-NOV-2000; 2000US-00721254.	
XX	PR	12-APR-2001; 2001US-00833328.	
XX	PA	(BAXT) BAXTER AG.	
XX	PI	Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;	
XX	PI	Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;	
XX	PI	Zimmermann K, Voelkel D;	
XX	DR	WPI; 2002-479950/51.	
XX	DR	N-PSDB; AAD39332.	
XX	PT	Novel isolated or substantially purified Von Willebrand factor-cleaving	
XX	PT	protease, useful for producing preparation for therapy of thrombosis and	
XX	PT	thromboembolic disease such as thrombotic thrombocytic purpura.	
XX	PS	Claim 1; Fig 5; 93pp; English.	
XX	CC	The invention relates to an isolated or substantially pure Von Willebrand	
XX	CC	factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for	
XX	CC	purifying vWF which involves providing vWF-cp as a ligand, contacting a	
XX	CC	solution comprising vWF with the polypeptide ligand under conditions	
XX	CC	where vWF is bound to the ligand and recovering from the ligand purified	
XX	CC	vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies	
XX	CC	which involves immunising an animal with vWF-cp and isolating the anti-	
XX	CC	vWF cp polypeptide antibodies from the animal. vWF-cp is useful for	
XX	CC	producing a preparation of prophylaxis and therapy of thrombosis and	
XX	CC	thromboembolic disease such as thrombotic thrombocytic purpura (TTP),	
XX	CC	Henoch-Schonlein purpura, preecclampsia, neonatal thrombocytopaenia or	
XX	CC	haemolytic-uraemic syndrome. vWF-cp can also be used for processing	
XX	CC	plasmatic or recombinantly produced vWF. The invention is useful for	
XX	CC	construction expression systems and generating transgenic animals which	
XX	CC	express the polypeptide in vivo. The present sequence is human vWF-cp	
XX	CC	protein. vWF-cp gene is located on chromosome 9	
XX	Sequence	1427 AA;	
Qy	Query Match	100.0%; Score 365; DB 5; Length 1427;	
Db	Best Local Similarity	100.0%; Pred. NO. 1.3e-30;	
	Matches	64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	WSSWGRSPCSRSCGGVTRRRQCNPRPFAFGGRACVGDADLAQEMCNTOACEKTQLEFM 60	
Db	387	WSSWGRSPCSRSCGGVTRRRQCNPRPFAFGGRACVGDADLAQEMCNTOACEKTQLEFM 446	
Qy	61	SOQC 64	
Db	447	SOQC 450	
RESULT 12			
ADD94038			
ID	ADD94038	standard; protein; 1427 AA.	
XX	AC	ADD94038;	
XX	XX	29-JAN-2004 (first entry)	
DT	XX	Human aggrecanase protein amino acid sequence.	
XX	XX	aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;	
XX	XX	cartilage degradation; osteoarthritis; inflammatory joint disease;	
KW	XX	antiarthritic; osteopathic; antiinflammatory;	
KW	XX	aggrecanase-associated disorder; osteoarthritis; inflammatory condition;	

KW	human; enzyme.	
XX	Homo sapiens.	
XX	US2003105313-A1.	
XX	05-JUN-2003.	
XX	25-JAN-2002; 2002US-00057487.	
XX	16-OCT-2001; 2001US-00978979.	
XX	(AMHP) AMERICAN HOME PROD CORP.	
XX	Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;	
XX	WPI; 2003-801251/75.	
XX	N-PSDB; ADD94037.	
XX	New isolated DNA molecule encoding an aggrecanase polypeptide for	
XX	producing a purified human aggrecanase protein which can be used to	
XX	develop inhibitors of aggrecanase.	
XX	Claim 14; SEQ ID NO 8; 24pp; English.	
XX	This invention relates to a novel human aggrecanase protein and the DNA	
XX	sequence which encodes it. Aggrecan is a major extracellular component of	
XX	articular cartilage. It is a proteoglycan responsible for providing	
XX	cartilage with its mechanical properties of compressability and	
XX	elasticity. A proteolytic activity (aggrecanase) is responsible for the	
XX	cleavage of aggrecan thereby having a role in cartilage degradation	
XX	associated with osteoarthritis and inflammatory joint disease. Compounds	
XX	which inhibit the activity of the protein of the invention may have	
XX	antiarthritic, osteopathic or antiinflammatory activity. The invention	
XX	may be used to produce a purified human aggrecanase protein. The protein	
XX	(or fragment) may be used to develop inhibitors of aggrecanase, using	
XX	three dimensional structural analysis or computer aided drug design. A	
XX	peptide which binds to aggrecanase is used to inhibit the proteolytic	
XX	degradation of aggrecan. The invention may be useful for the development	
XX	of therapeutics for the treatment of aggrecanase-associated disorders	
XX	such as, osteoarthritis and other inflammatory conditions. The present	
XX	sequence is that of the human aggrecanase protein (full length sequence)	
XX	of the invention.	
XX	Sequence 1427 AA;	
Qy	Query Match	100.0%; Score 365; DB 7; Length 1427;
Db	Best Local Similarity	100.0%; Pred. NO. 1.3e-30;
	Matches	64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1	WSSWGRSPCSRSCGGVTRRRQCNPRPFAFGGRACVGDADLAQEMCNTOACEKTQLEFM 60
Db	387	WSSWGRSPCSRSCGGVTRRRQCNPRPFAFGGRACVGDADLAQEMCNTOACEKTQLEFM 446
Qy	61	SOQC 64
Db	447	SOQC 450
RESULT 13		
ABB98125		
ID	ABB98125	standard; protein; 1445 AA.
XX	AC	ABB98125;
XX	XX	17-OCT-2002 (first entry)
DT	XX	Human PMMM Incyte ID 7473607CD1.
XX	XX	Human; PMMM; protein modification and maintenance molecule;
XX	XX	anticonvulsant; neuroprotective; nootropic; cytostatic; antipsoriatic;
KW	XX	antisthmatic; dermatological; antidiabetic; antiparkinsonian;
KW	XX	antianaemic; antiinflammatory; antiulcer; antilanginal; cardiant;

KW hepatotropic; osteopathic; antiemetic; antipyrretic; virucide;
 KW antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
 KW haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
 KW cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
 KW antitumour; anirheumatic; immunosuppressive; anti-allergic; antithyroid;
 KW rephrotropic; anti-gout; thymometric; antiarthritic; uropathic;
 KW ophthalmological; antiparasitic; tranquiliser; vulnerary; keratolytic;
 KW auditory; antiseborrheic; antidepressant; neuroleptic; antinfertility;
 KW anthelmintic; protozoacide; Crohn's disease; hypertension; autoimmune;
 KW inflammatory; anaemic; cell proliferative; developmental; epithelial;
 KW scabies; neurological; Alzheimer's disease; reproductive;
 KW ectopic pregnancy; gene therapy; vaccine; disorder;
 KW procollagen I N-proteinase.
 XX
 XX Homo sapiens.
 XX
 XX WO200246383-A2.
 XX
 XX 13-JUN-2002.
 XX
 XX 05-DEC-2001; 2001WO-US046964.
 XX
 XX 08-DEC-2000; 2000US-0254399P.
 XX 21-DEC-2000; 2000US-0257803P.
 XX 05-JAN-2001; 2001US-0360110P.
 XX 19-JAN-2001; 2001US-0362851P.
 XX 25-JAN-2001; 2001US-0364623P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Yue H, Azimzai Y, Kallick DA, Baughn MR, Griffin JA, Swarnakar A;
 XX Lal PG, Wallia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;
 XX Ramkumar J, Thangavelu K, Lu Y, Warren BA, Lu DAM, Lee EA;
 XX Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;
 XX Sanjanwala MM;
 XX
 XX WPI; 2002-519664/55.
 XX N-PSDB; ABQ75946.
 XX
 XX New isolated Protein Modification and Maintenance polypeptides, useful
 XX for diagnosis, and treatment of e.g. gastrointestinal disorders.
 XX
 XX Claim 1 (a); Page 154-157; 200pp; English.
 XX
 XX The invention relates to an isolated Protein Modification and Maintenance
 XX (PMW) polypeptide. Polypeptides of the invention may be used in the
 XX diagnosis, treatment and prevention of disorders associated with
 XX decreased expression or activity of PMW. These include gastrointestinal
 XX disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
 XX hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
 XX proliferative disorders, developmental disorders, epithelial disorders
 XX (e.g. scabies), neurological disorders (e.g. Alzheimer's disease)
 XX reproductive disorders (e.g. ectopic pregnancy), and in gene therapy or a
 XX vaccine for such diseases. They may also be used in the assessment of the
 XX effects of exogenous compound on the expression of nucleic acid and amino
 XX acid sequences of protein modification and maintenance molecules. The
 XX current sequence represents a human PMW of the invention, which has been
 XX found to have homology with B. taurus procollagen I N-proteinase
 XX
 XX Sequence 1445 AA;
 XX
 XX Query Match 100.0%; Score 365; DB 5; Length 1445;
 XX Best Local Similarity 100.0%; Pred. No. 1.3e-30;
 XX Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0
 QY 1 WSSWGPRSPCRSGGGVYVTRRRQCNPNRPAFGGRACVADLQAEKMTQACEKTLQLEFM 60
 Db 390 WSSWGPRSPCRSGGGVYVTRRRQCNPNRPAFGGRACVADLQAEKMTQACEKTLQLEFM 449
 QY 61 SQQC 64
 Db 450 SQQC 453

```
XX Homo sapiens.
OS
XX
XX WO200159133-A1.
XX
XX 16-AUG-2001.
XX
XX 12-FEB-2001; 2001WO-EP001525.
XX
XX 14-FEB-2000; 2000EP-00102955.
XX
XX (MERE ) MERCK PATENT GMBH.
XX
XX Duecker K, Hentsch B, Hoheisel J, Frohne M;
XX
XX WPI; 2001-483569/52.
XX
XX N-PSDB; AAH47048.
XX
XX Larynx carcinoma associated protein-1 polypeptide for the treatment of
XX carcinomas, metastasis, arthritis, osteoporosis, immune disorders,
XX stroke, ischemia, autoimmune disease, angiogenesis, skin disorders and
XX organ malfunctions.
XX
XX Claim 1; Page 33-39; 47pp; English.
XX
XX This represents a human Larynx carcinoma associated protein-1 (LarCAP-1)
XX polypeptide. The LarCAP-1 polypeptide can be expressed by standard
XX recombinant methodology. The LarCAP-1 polypeptide and polynucleotide are
XX useful in diagnostic assays and for the treatment of carcinomas,
XX metastasis, arthritis, osteoporosis, immune disorders, stroke, ischemia,
XX autoimmune disease, angiogenesis, skin disorders and organ malfunctions
XX especially heart hypertrophy
XX
XX Sequence 1048 AA;
XX
XX Query Match 56.4%; Score 206; DB 4; Length 1048;
XX Best Local Similarity 53.1%; Pred. No. 1.3e-13;
XX Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;
XX
XX QY 1 WSSWGPRSPCSRSGGQVTRRRQCNPRPAFGGRACVGDQAEMCNTOACEKTQLEFM 60
XX |||||
XX 344 WSSWTKFGSCSRSGGVRSRSCNNPSPAYGGRCLGPMFYYQVCNSECPCGTIEDFR 403
XX |||||
XX QY 61 SQQC 64
XX :|||
XX Db 404 AQQC 407
```

Search completed: March 13, 2004, 07:39:06
Job time : 10.3667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 5.06667 Seconds
(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473

Perfect score: 365
Sequence: 1 WSWGFRPCSRSGGVVT.....EMNTQACEKTQLEFMGQQC 64

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pcp.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pcp.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pcp.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pcp.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pcp.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pcp.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pcp.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pcp.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pcp.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pcp.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pcp.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pcp.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pcp.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pcp.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pcp.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pcp.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pcp.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	365	100.0	365	9	US-09-778-979-5
2	365	100.0	365	14	US-10-057-487-5
3	365	100.0	842	14	US-10-222-334-4
4	365	100.0	1416	9	US-09-836-712-2
5	365	100.0	1427	14	US-10-222-334-2
6	365	100.0	1427	14	US-10-057-487-8
7	206	56.4	1156	13	US-10-014-070-5
8	206	56.4	1189	16	US-10-311-035-10
9	206	56.4	1213	9	US-09-938-330-14
10	206	56.4	1216	9	US-09-938-330-12
11	206	56.4	1222	9	US-09-938-330-10
12	206	56.4	1229	9	US-09-938-330-8
13	206	56.4	1223	13	US-10-014-070-2
14	206	56.4	1232	14	US-10-205-368-4
15	206	56.4	1232	9	US-09-938-330-18

15	206	56.4	1233	14	US-10-205-368-2
16	206	56.4	1235	9	US-09-938-330-16
17	206	56.4	1249	9	US-09-938-330-22
18	206	56.4	1252	9	US-09-938-330-20
19	206	56.4	1360	14	US-10-205-368-5
20	206	56.4	1360	14	US-10-205-368-5
21	195	53.4	791	11	US-09-981-151A-4
22	195	53.4	791	11	US-09-981-151A-4
23	192	52.6	952	11	US-09-981-151A-8
24	192	52.6	1224	13	US-10-217-774-4
25	192	52.6	1224	13	US-10-217-774-4
26	191	52.3	1365	15	US-10-296-616-2
27	191	52.3	1365	15	US-10-296-616-2
28	191	52.3	1365	15	US-10-296-616-2
29	191	52.3	1365	15	US-10-296-616-2
30	191	52.3	1365	15	US-10-296-616-2
31	190	52.1	862	14	US-10-274-639-10
32	190	52.1	862	14	US-10-274-639-10
33	190	52.1	862	14	US-10-274-639-10
34	190	52.1	862	14	US-10-274-639-10
35	189	51.8	1081	15	US-10-161-493-124
36	187	51.2	924	15	US-10-161-493-124
37	185	50.7	1205	14	US-10-161-493-124
38	185	50.7	1205	14	US-10-161-493-124
39	183	50.1	989	9	US-10-295-027-528
40	183	50.1	989	9	US-10-295-027-528
41	183	50.1	989	9	US-10-295-027-528
42	179	49.0	874	9	US-09-321-987B-5
43	177	48.5	481	13	US-09-918-171A-15
44	177	48.5	481	13	US-09-918-171A-15
45	177	48.5	481	13	US-09-918-171A-15

ALIGNMENTS

RESULT 1
US-09-978-979-5
; Sequence 5, Application US/09978979
; Patent No. US20020151702A1
; GENERAL INFORMATION:
; APPLICANT: Racie, Lisa, A.
; Twine, Natalie, C.
; Agostino, Michael, J.
; Wolfman, Neil
; Morris, Elisabeth
TITLE OF INVENTION: Aggracanthase Molecules
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Home Products Corporation
STREET: One Campus Drive
CITY: Parsippany
STATE: New Jersey
COUNTRY: USA
ZIP: 07054
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/978,979
FILING DATE: 16-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/60/241,469
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Kapinos, Ellen J.
REGISTRATION NUMBER: 32,345
REFERENCE/DOCKET NUMBER: GI 5435p
TELECOMMUNICATION INFORMATION:
TELEPHONE: (973) 660-5000
TELEFAX: (973) 683-4117

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 365 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-978-979-5
Query Match 100.0%; Score 365; DB 9; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-32;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 60
Db 59 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 118
Qy 61 SQQC 64
Db 119 SQQC 122
RESULT 2
US-10-057-487-5
Sequence 5, Application US/10057487
Publication No. US20030105313A1
GENERAL INFORMATION:
APPLICANT: Wyeth
TITLE OF INVENTION: Aggreganase Molecules
FILE REFERENCE: 08702.0073
CURRENT APPLICATION NUMBER: US/10/057,487
CURRENT FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: 60/241,469
PRIOR FILING DATE: 2000-10-18
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
TYPE: PRT
ORGANISM: homo sapiens
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: unknown amino acid
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (365)..(365)
OTHER INFORMATION: unknown amino acid
US-10-057-487-5
Query Match 100.0%; Score 365; DB 14; Length 365;
Best Local Similarity 100.0%; Pred. No. 2.8e-32;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 60
Db 59 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 118
Qy 61 SQQC 64
Db 119 SQQC 122
RESULT 3
US-10-222-334-4
Sequence 4, Application US/10222334
Publication No. US20030073116A1
GENERAL INFORMATION:
APPLICANT: Ginsburg, David
APPLICANT: Levy, Gallia
APPLICANT: Tsai, Han-Mou
TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
FILE REFERENCE: UM-07288

CURRENT APPLICATION NUMBER: US/10/222,334
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 60/312,834
PRIOR FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 842
TYPE: PRT
ORGANISM: Homo sapiens
US-10-222-334-4
Query Match 100.0%; Score 365; DB 14; Length 842;
Best Local Similarity 100.0%; Pred. No. 6.3e-32;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 60
Db 387 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 446
Qy 61 SQQC 64
Db 447 SQQC 450
RESULT 4
US-09-836-712-2
Sequence 2, Application US/09836712
Patent No. US20010049106A1
GENERAL INFORMATION:
APPLICANT: PFIZER INC.
TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
FILE REFERENCE: FC10851A
CURRENT APPLICATION NUMBER: US/09/836,712
CURRENT FILING DATE: 2001-04-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1416
TYPE: PRT
ORGANISM: Human
US-09-836-712-2
Query Match 100.0%; Score 365; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1e-31;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 60
Db 410 WSSWGRSPCSRSCGGVVTTRRQCNPRPFAFGGRACVADLAEMCNTOACEKTQLEFM 469
Qy 61 SQQC 64
Db 470 SQQC 473
RESULT 5
US-10-222-334-2
Sequence 2, Application US/10222334
Publication No. US20030073116A1
GENERAL INFORMATION:
APPLICANT: Ginsburg, David
APPLICANT: Levy, Gallia
APPLICANT: Tsai, Han-Mou
TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
FILE REFERENCE: UM-07288
CURRENT APPLICATION NUMBER: US/10/222,334
CURRENT FILING DATE: 2002-08-16
PRIOR APPLICATION NUMBER: 60/312,834
PRIOR FILING DATE: 2001-08-16
NUMBER OF SEQ ID NOS: 78
SOFTWARE: PatentIn version 3.1


```
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2

Query Match      100.0%; Score 365; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1e-31;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 60
Db 387 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 446

Qy 61 SQQC 64
Db 447 SQQC 450

RESULT 6
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggreganase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-057-487-8

Query Match      100.0%; Score 365; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1e-31;
Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 60
Db 387 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 446

Qy 61 SQQC 64
Db 447 SQQC 450
```

```
; ORGANISM: Homo sapiens
US-10-014-070-5

Query Match      56.4%; Score 206; DB 13; Length 1156;
Best Local Similarity 53.1%; Pred. No. 2.6e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 60
Db 488 WSSWTKFGSCRSRSCGGVRSRSCNNPSPAYGRCPLGPMFEYQVCNSECPTGYEDFR 547

Qy 51 SQQC 64
Db 548 AQQC 551

RESULT 8
US-10-311-035-10
; Sequence 10, Application US/10311035
; Publication No. US20040023243A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: ELLIOTT, Vicki
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LAL, Preeti
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: LEE, Ernestine A.
; APPLICANT: HAFALIA, April
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: TANG, Y. Tom
; APPLICANT: WALSH, Roderick T.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: LU, Yan
; APPLICANT: RAMKUMAR, Jayalaximi
; APPLICANT: XU, Yuming
; APPLICANT: REDDY, Roopa
; APPLICANT: DAS, Depopriya
; APPLICANT: KEARNEY, Liam
; APPLICANT: KALLICK, Deborah A.
; TITLE OF INVENTION: Proteases
; FILE REFERENCE: PI-0123 PCT
; CURRENT APPLICATION NUMBER: US/10/311,035
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: 60/212,336; 60/213,955; 60/215,396; 60/216,821; 60/218,941
; PRIOR FILING DATE: 2000-06-16; 2000-06-22; 2000-06-29; 2000-07-07; 2000-07-14
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023243A1 7477386CD1
US-10-311-035-10

Query Match      56.4%; Score 206; DB 16; Length 1189;
Best Local Similarity 53.1%; Pred. No. 2.7e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 WSSWGPRSPCSRSCGGVVTRRRQCNPRPAFGGRACVGVADLQAEMLNTQACEKTQLEFM 60
Db 558 WSSWTKFGSCRSRSCGGVRSRSCNNPSPAYGRCPLGPMFEYQVCNSECPTGYEDFR 617
```

Best Local Similarity 53.1%; Pred. No. 2.8e-14; Mismatches 10; Indels 0; Gaps 0; Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 WSSWGRSPCRSCGGVVTTRRRQCNPRPAGFGRACVADLQAEKQLEFM 60
Db 555 WSSWTKFGSCRSRSCGGVRSRSCNNPSPAYGGRPCLGPMFEYQVCNSECPCGTVEDFR 614

Qy 61 SQQC 64
Db 615 AQQC 618

RESULT 11
US-09-938-330-10
; Sequence 10, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1219
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-10

Query Match 56.4%; Score 206; DB 9; Length 1219;
Best Local Similarity 53.1%; Pred. No. 2.8e-14; Mismatches 10; Indels 0; Gaps 0; Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 WSSWGRSPCRSCGGVVTTRRRQCNPRPAGFGRACVADLQAEKQLEFM 60
Db 555 WSSWTKFGSCRSRSCGGVRSRSCNNPSPAYGGRPCLGPMFEYQVCNSECPCGTVEDFR 614

Qy 61 SQQC 64
Db 615 AQQC 618

RESULT 12
US-09-938-330-8
; Sequence 8, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0

Qy 61 SQQC 64
Db 618 AQQC 621

RESULT 9
US-09-938-330-14
; Sequence 14, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 1213
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-14

Query Match 56.4%; Score 206; DB 9; Length 1213;
Best Local Similarity 53.1%; Pred. No. 2.8e-14; Mismatches 10; Indels 0; Gaps 0; Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

Qy 1 WSSWGRSPCRSCGGVVTTRRRQCNPRPAGFGRACVADLQAEKQLEFM 60
Db 555 WSSWTKFGSCRSRSCGGVRSRSCNNPSPAYGGRPCLGPMFEYQVCNSECPCGTVEDFR 614

Qy 61 SQQC 64
Db 615 AQQC 618

RESULT 10
US-09-938-330-12
; Sequence 12, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encodin
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 1216
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-12

Query Match 56.4%; Score 206; DB 9; Length 1216;

```
; SEQ ID NO 8
; LENGTH: 1222
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-8

Query Match
  56.4%; Score 206; DB 9; Length 1222;
Best Local Similarity 53.1%; Pred. No. 2.8e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGRSPGSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
Db 555 WSWTFGSCSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
QY 61 SQQC 64
Db 615 AQQC 618
```

```
RESULT 13
US-10-014-070-2
; Sequence 2, Application US/10014070
; Publication No. US20020119555A1
; GENERAL INFORMATION:
; APPLICANT: Bandaru, Rajasekhar
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Spurling, Heidi Lynn
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 53014, A Human Metalloprotease Family
; TITLE OF INVENTION: Member and Uses Therefor
; FILE REFERENCE: MPI2000-523P1RCP1(M)
; CURRENT APPLICATION NUMBER: US/10/014,070
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: 60/258,373
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-014-070-2
```

```
Query Match
  56.4%; Score 206; DB 13; Length 1223;
Best Local Similarity 53.1%; Pred. No. 2.8e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGRSPGSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
Db 555 WSWTFGSCSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
QY 61 SQQC 64
Db 615 AQQC 618
```

```
RESULT 14
US-10-205-368-4
; Sequence 4, Application US/10205368
; Publication No. US20030092621A1
; GENERAL INFORMATION:
; APPLICANT: AGOSTINO, MICHAEL J.
; APPLICANT: RACIE, LISA A.
; TITLE OF INVENTION: AGGRECANASE MOLECULES
; FILE REFERENCE: 08702.0094-00000
; CURRENT APPLICATION NUMBER: US/10/205,368
; CURRENT FILING DATE: 2002-07-26
; PRIOR APPLICATION NUMBER: 60/308,386
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: 60/315,887
; PRIOR FILING DATE: 2001-08-29
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin Ver. 2.1
```

```
; SEQ ID NO 4
; LENGTH: 1223
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-368-4

Query Match
  56.4%; Score 206; DB 14; Length 1223;
Best Local Similarity 53.1%; Pred. No. 2.8e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGRSPGSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
Db 555 WSWTFGSCSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
QY 61 SQQC 64
Db 615 AQQC 618
```

```
RESULT 15
US-09-938-330-18
; Sequence 18, Application US/09938330
; Patent No. US20020115838A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hilbun, Erin
; APPLICANT: Scoville, John
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hu, Yi
; APPLICANT: Turner, C. Alexander Jr.
; TITLE OF INVENTION: No. US20020115838A1 Human Proteases and Polynucleotides Encod
; FILE REFERENCE: LEX-0237-USA
; CURRENT APPLICATION NUMBER: US/09/938,330
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: US 60/227,104
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: US 60/233,796
; PRIOR FILING DATE: 2000-09-19
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 1232
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-938-330-18
```

```
Query Match
  56.4%; Score 206; DB 9; Length 1232;
Best Local Similarity 53.1%; Pred. No. 2.8e-14;
Matches 34; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGRSPGSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
Db 555 WSWTFGSCSRSCGGGVVTRRRQCNPRPAFGGRACVADLQAEQMCNTQACEKTLQLEFM 60
QY 61 SQQC 64
Db 615 AQQC 618
```

Search completed: March 13, 2004, 08:07:05
Job time : 5.06667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:34:04 ; Search time 2.7 Seconds
(without alignments)
1223.727 Million cell updates/sec

Title: US-09-836-712-2_COPY_410_473
Perfect score: 365
Sequence: 1 WSSWGPRSPCRSCGGVVT.....EMCMTQACEKTOLEFMSQQC 64

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	192	52.6	1224	4	US-09-930-872-4
2	191	52.3	1882	4	US-09-369-364A-13
3	183	50.1	1205	4	US-09-491-522-11
4	179	49.0	874	4	US-09-369-364A-15
5	177	48.5	481	4	US-09-130-491-8
6	177	48.5	1211	4	US-09-491-522-5
7	175	47.9	905	4	US-09-369-364A-9
8	165	45.2	1081	4	US-09-369-364A-17
9	163	44.7	997	4	US-09-369-364A-7
10	162	44.4	1172	1	US-08-313-288B-19
11	161	44.1	930	4	US-09-369-364A-2
12	161	44.1	1170	1	US-08-313-288B-20
13	157	43.0	239	5	PCT-US93-01552-1
14	155.5	42.6	551	4	US-09-130-491-16
15	155.5	42.6	608	4	US-09-130-491-13
16	155.5	42.6	727	4	US-09-445-023A-1
17	155.5	42.6	727	4	US-09-445-023A-12
18	155.5	42.6	949	4	US-09-568-559-2
19	155.5	42.6	967	4	US-09-130-491-2
20	154.5	42.3	757	4	US-09-963-791-24
21	154.5	42.3	837	4	US-09-122-126B-2
22	154.5	42.3	837	4	US-09-634-286A-2
23	154.5	42.3	908	4	US-09-963-791-2
24	153	41.9	518	4	US-09-369-364A-22
25	153	41.9	930	4	US-09-122-126B-15
26	153	41.9	930	4	US-09-634-286A-15
27	150.5	41.2	2165	4	US-09-800-729-155

Sequence 12, Appl
Sequence 12, Appl
Patent No. 5426100
Sequence 15, Appl
Sequence 3, Appl
Sequence 5, Appl
Sequence 7, Appl
Sequence 21, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 3, Appl
Sequence 1, Appl
Sequence 125, Appl
Sequence 89, Appl
Sequence 18, Appl
Sequence 5, Appl
Sequence 5, Appl

1 US-07-646-531D-12
2 US-08-488-273-12
6 5426100-12
469 1 US-08-313-288B-15
441 3 US-08-985-526-3
859 4 US-09-369-364A-5
943 2 US-08-808-982-7
943 3 US-09-306-902A-7
525 4 US-09-369-364A-21
1156 3 US-08-996-083-1
1156 3 US-09-429-516-1
1156 3 US-09-429-516-3
218 3 US-08-585-526-1
262 4 US-09-800-729-125
1745 4 US-09-800-729-89
59 4 US-09-197-770B-18
898 2 US-08-808-982-5
898 3 US-09-306-902A-5

ALIGNMENTS

RESULT 1
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fiddle, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 6448388el Human Proteases and Polynucleotides Encoding the
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match 52.6%; Score 192; DB 4; Length 1224;
Best Local Similarity 46.9%; Pred. No. 5.7e-13;
Matches 30; Conservative 12; Mismatches 22; Indels 0; Gaps 0;
QY 1 WSSWGPRSPCRSCGGVVTERRRCNNPRPAPGACVADLQAEKMTQACEKTOLEFM 60
Db 589 WSDWSWSPCRSCGGVSHRSLCTNPKPSHGKFCGEGTETLKLNSQKCPDSDVDFR 648
QY 61 SQQC 64
Db 649 AAQC 652

RESULT 2
US-09-369-364A-13
; Sequence 13, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apté, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 1882

```

; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = C
; NAME/KEY: MOD RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Y
US-09-369-364A-13

Query Match      52.3%; Score 191; DB 4; Length 1882;
Best Local Similarity 46.9%; Pred. No. 1.1e-12;
Matches 30; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

Qy 1 WSSWGRPRPCSRSCGGVTRRRQCNPRPAFGGRACVGDLOAEMCNTQACEKTQLEFM 60
Db 538 WGSWSFPGTCRTCGGKIKTAIRECNRPBKNGKVCVGRMRKFKSCNTPCLKQKQKDR 597
Qy 61 SQQC 64
Db 598 DEQC 601

RESULT 3
US-09-491-522-11
; Sequence 11; Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapierre, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE, AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds, LLP
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B.
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 850-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-11

Query Match      50.1%; Score 183; DB 4; Length 1205;
Best Local Similarity 46.9%; Pred. No. 5.4e-12;
Matches 30; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = C
; NAME/KEY: MOD RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Y
US-09-369-364A-15

Query Match      49.0%; Score 179; DB 4; Length 874;
Best Local Similarity 43.8%; Pred. No. 1.1e-11;
Matches 28; Conservative 11; Mismatches 25; Indels 0; Gaps 0;

Qy 1 WSSWGRPRPCSRSCGGVTRRRQCNPRPAFGGRACVGDLOAEMCNTQACEKTQLEFM 60
Db 430 WGSWSHFGTCRTCGGKIKTAIRECNRPBKNGKVCVGRMRKFKSCNTPCKMKQKDR 489
Qy 61 SQQC 64
Db 490 EQQC 493

RESULT 5
US-09-130-491-8
; Sequence 8; Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-8

Query Match      48.5%; Score 177; DB 4; Length 481;
Best Local Similarity 43.8%; Pred. No. 1.1e-11;
Matches 28; Conservative 9; Mismatches 27; Indels 0; Gaps 0;

Qy 1 WSSWGRPRPCSRSCGGVTRRRQCNPRPAFGGRACVGDLOAEMCNTQACEKTQLEFM 60
Db 121 WGPWGPWQGSRTCGGGIQFSNRECDNPAPONGGRFCLGERVKYQSKTEECPPNGKSR 180

```

QY 61 SQQC 64
Db 181 EQQC 184

RESULT 6
US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Collige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE.
; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1211 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-09-491-522-5

Query Match 48.5%; Score 177; DB 4; Length 1211;
Best Local Similarity 45.3%; Pred. No. 2.5e-11;
Matches 29; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRCGGGVVTRRQCNNPRPAFGGRACVGDADLQAEIMCMTQACEKTLQEFM 60
Db 564 WCAWSPFGSCRTCGTGKFRTRQCDNPHFANGGRTCSGLAYDFQLCSRQDCPSLADFR 623

QY 61 SQQC 64
Db 624 EQQC 627

RESULT 7
US-09-369-364A-9
; Sequence 9, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi

; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 905
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-8
US-09-369-364A-9

Query Match 47.9%; Score 175; DB 4; Length 905;
Best Local Similarity 43.8%; Pred. No. 3.1e-11;
Matches 28; Conservative 9; Mismatches 27; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRCGGGVVTRRQCNNPRPAFGGRACVGDADLQAEIMCMTQACEKTLQEFM 60
Db 545 WGPWRPWGQCSTCGGGIQFSNRECDNPMQNGRFLGERVKYQSCNTECPNGKSPR 604

QY 61 SQQC 64
Db 605 EQQC 608

RESULT 8
US-09-369-364A-17
; Sequence 17, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-10
US-09-369-364A-17

Query Match 45.2%; Score 165; DB 4; Length 1081;
Best Local Similarity 42.2%; Pred. No. 4.5e-10;
Matches 27; Conservative 8; Mismatches 29; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRCGGGVVTRRQCNNPRPAFGGRACVGDADLQAEIMCMTQACEKTLQEFM 60
Db 527 WGPWRPWGQCSTCGGGVSSSRHCDSPRPTIGGKCYLGERRRHRSNCNTDCCPPGSDFR 586

QY 61 SQQC 64
Db 587 EQQC 590

RESULT 9
US-09-369-364A-7
; Sequence 7, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7

LENGTH: 997
TYPE: PRT
ORGANISM: Homo sapiens ADAMTS-7
US-09-369-364A-7

Query Match 44.7%; Score 163; DB 4; Length 997;
Best Local Similarity 46.9%; Pred. No. 6.9e-10;
Matches 30; Conservative 5; Mismatches 29; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVWTRRRQNNPRPAFGGRACVADLAQEMCNTOACEKTOLEFM 60
DB 541 WSGWASWISCSRCGNGVQSAERQCTPTPKYKGRYCVGERKFRCLNQLQACPAGRSFR 600

QY 61 SQQC 64
DB 601 HVQC 604

RESULT 10
US-08-313-288B-19
Sequence 19, Application US/08313288B
Patent No. 5750502

GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:

INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-19

Query Match 44.4%; Score 162; DB 1; Length 1172;
Best Local Similarity 48.1%; Pred. No. 1e-09;
Matches 25; Conservative 10; Mismatches 17; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVWTRRRQNNPRPAFGGRACVADLAQEMCNTOAC 52
DB 497 WSPWFSWACTVTCAGGIRNTRVCNPFEPYGGKACVGDYQERQCNKRSC 548

RESULT 11
US-09-369-364A-2
Sequence 2, Application US/09369364A
Patent No. 6391610

Query Match 44.4%; Score 161; DB 1; Length 1170;
Best Local Similarity 51.9%; Pred. No. 1.3e-09;
Matches 29; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVWTRRRQNNPRPAFGGRACVADLAQEMCNTOACEKTOLEFM 60
DB 570 WSGWGPWGQCSRSRCGGVQFAYRHCCNPPAPRNSGRYCTGKRAIYRSCVTPCPPNGKSF 629

GENERAL INFORMATION:
APPLICANT: Apte, Suneel
APPLICANT: Hurskainen, Riina L.
APPLICANT: Hirshata, Satoshi
TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
FILE REFERENCE: 26473/4007/10-30-00
CURRENT APPLICATION NUMBER: US/09/369,364A
CURRENT FILING DATE: 1999-08-06
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 930
TYPE: PRT
ORGANISM: mus musculus ADAMTS-5
US-09-369-364A-2

Query Match 44.1%; Score 161; DB 4; Length 930;
Best Local Similarity 45.3%; Pred. No. 1.1e-09;
Matches 29; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVWTRRRQNNPRPAFGGRACVADLAQEMCNTOACEKTOLEFM 60
DB 570 WSGWGPWGQCSRSRCGGVQFAYRHCCNPPAPRNSGRYCTGKRAIYRSCVTPCPPNGKSF 629

QY 61 SQQC 64
DB 630 HEQC 633

RESULT 12
US-08-313-288B-20
Sequence 20, Application US/08313288B
Patent No. 5750502

GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-20

Query Match 44.1%; Score 161; DB 1; Length 1170;
Best Local Similarity 51.9%; Pred. No. 1.3e-09;
Matches 29; Conservative 2; Mismatches 33; Indels 0; Gaps 0;

QY 1 WSSWGPSPCSRCGGVWTRRRQNNPRPAFGGRACVADLAQEMCNTOACEKTOLEFM 60
DB 570 WSGWGPWGQCSRSRCGGVQFAYRHCCNPPAPRNSGRYCTGKRAIYRSCVTPCPPNGKSF 629

QY 61 SQQC 64
DB 630 HEQC 633

RESULT 13
US-08-313-288B-20
Sequence 20, Application US/08313288B
Patent No. 5750502

GENERAL INFORMATION:
APPLICANT: Jessell, Thomas M. and Avihu Klar
TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A
NOVEL SECRETED PROTEIN, F-SPONDIN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/313,288B
FILING DATE: January 5, 1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
TELEX:

INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 1170 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-313-288B-20

Matches 27; Conservative 5; Mismatches 20; Indels 0; Gaps 0;

QY 1 WSSWGRSPCSRSCGGVTVTRRQCNNRPAFGGRACVGDADLAEMCNTOAC 52
Db 495 WGPWSPWDICSVTCGGGVQKSRSLCNPPTQFGGKDCVGDVTENQICNKQDC 546

RESULT 13
PCT-US93-01652-1
; Sequence 1, Application PC/TUS9301652
; GENERAL INFORMATION:
; APPLICANT: Bouck, Noel P.
; APPLICANT: Polverini, Peter J.
; APPLICANT: Good, Deborah J.
; APPLICANT: Frazier, William A.
; TITLE OF INVENTION: Method and Composition for
; TITLE OF INVENTION: Inhibiting Angiogenesis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSES: Tilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01652
; FILING DATE: 19930222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,656
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/464,369
; FILING DATE: 12-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92005-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US93-01652-1

Query Match 43.0%; Score 157; DB 5; Length 239;
Best Local Similarity 51.9%; Pred. No. 7.7e-10;
Matches 27; Conservative 4; Mismatches 21; Indels 0; Gaps 0;
QY 1 WSSWGRSPCSRSCGGVTVTRRQCNNRPAFGGRACVGDADLAEMCNTOAC 52
Db 184 WGPWSPWDICSVTCGGGVQKSRSLCNPPTQFGGKDCVGDVTENQICNKQDC 235

RESULT 14
US-09-130-491-16
; Sequence 16, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83

; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 551
; TYPE: PRT
; ORGANISM: Rattus rattus
US-09-130-491-16

Query Match 42.6%; Score 155.5; DB 4; Length 551;
Best Local Similarity 41.5%; Pred. No. 2.5e-09;
Matches 27; Conservative 9; Mismatches 28; Indels 1; Gaps 1;
QY 1 WSSWGRSPCSRSCGGVTVTRRQCNNRPAFGGRACVGDADLAEMCNTOAC-EKTQLEF 59
Db 146 WGPWSPWDICSVTCGGGVQYTMRECDNPVKNNGKYCEGKRVYRSCNIEDCPDNNKTF 205
QY 60 MSQQC 64
Db 206 REEQC 210

RESULT 15
US-09-130-491-13
; Sequence 13, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Goodearl, Douglas A.
; APPLICANT: Holtzman, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 608
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-130-491-13

Query Match 42.6%; Score 155.5; DB 4; Length 608;
Best Local Similarity 41.5%; Pred. No. 2.8e-09;
Matches 27; Conservative 9; Mismatches 28; Indels 1; Gaps 1;
QY 1 WSSWGRSPCSRSCGGVTVTRRQCNNRPAFGGRACVGDADLAEMCNTOAC-EKTQLEF 59
Db 257 WGPWSPWDICSVTCGGGVQYTMRECDNPVKNNGKYCEGKRVYRSCNIEDCPDNNKTF 316
QY 60 MSQQC 64
Db 317 REEQC 321

Search completed: March 13, 2004, 07:46:08
Job time : 2.7 secs


```
A:Gene: CESP:F25H8.3
A:Map position: 4
A:Introns: 31/1, 52/1, 135/2, 193/3, 216/1, 266/1, 495/2, 547/3, 584/3, 634/2, 744/1, 811/1
Query Match 32.4%; Score 112; DB 2; Length 2165;
Best Local Similarity 38.2%; Pred. No. 0.00012;
Matches 21; Conservative 6; Mismatches 26; Indels 2; Gaps 1;
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
DB 1371 WATGPWTACATCGNGIQRLKLC--RDHVRDLDPDEYCNHLDKXVSTNCRLRDC 1423
RESULT 3
T22545
Hypothetical protein F53B6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T22545
R:White, S.
Submitted to the EMBL Data Library, October 1996
A:Reference number: Z19578
A:Accession: T22545
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1059 <WIL>
A:Cross-references: EMBL:Z81086; PIDN:CAB03121.1; GSPDB:GN00019; CESP:F53B6.2
A:Experimental source: clone F53B6
C:Genetics:
A:Gene: CESP:F53B6.2
A:Map position: 1
A:Introns: 38/3; 92/3; 131/3; 169/1; 236/2; 282/3; 349/2; 376/2; 420/1; 576/3; 759/3; 841/1
Query Match 29.3%; Score 101.5; DB 2; Length 1059;
Best Local Similarity 33.9%; Pred. No. 0.001;
Matches 19; Conservative 7; Mismatches 21; Indels 9; Gaps 2;
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLKPVTVRGCGWAG-PC 55
DB 296 WEAGKWSKCTASCGGVRRHVACVG-----GSDCDEGGRPRQETTCYAGIPC 343
RESULT 4
T18517
Procollagen N-endopeptidase (EC 3.4.24.14) I - bovine
N:Alternate names: procollagen N-proteinase
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T18517
R:Collge, A.; Nugsens, B.V.; Lapiere, C.M.
Submitted to the EMBL Data Library, February 1996
A:Description: Cloning of the cDNA of the bovine procollagen I N-proteinase.
A:Reference number: Z18941
A:Accession: T18517
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1205 <COL>
A:Cross-references: EMBL:X96389; NID:e990769; PID:e228215; PIDN:CRA65253.1
A:Experimental source: skin
C:Genetics:
A:Gene: PC I-NP
C:Function:
A:Description: catalyzes cleavage of the propeptides of type I and II collagens prior to
C:Keywords: hydrolase, metalloproteinase
Query Match 29.3%; Score 101.5; DB 2; Length 1205;
Best Local Similarity 38.2%; Pred. No. 0.0011;
Matches 21; Conservative 6; Mismatches 23; Indels 5; Gaps 2;
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
DB 973 WRAGSWSQSVTCGNGTQRPVLC----RTADDSFGVCRE-ERPETARICRLGFC 1022
```

RESULT 5

```
C89114
Protein C37C3.6a [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89114
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:title: Genome sequence of the nematode C. elegans: a platform for investigating biol
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C.e
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; a
A:Accession: C89114
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1558 <STO>
A:Cross-references: GB:chr_V; PIDN:AAC25867.1; PID:g3294501; GSPDB:GN00023; CESP:C37C3
C:Genetics:
A:Gene: C37C3.6a
A:Map position: 5
```

```
Query Match 27.5%; Score 95; DB 2; Length 1558;
Best Local Similarity 34.5%; Pred. No. 0.0079;
Matches 19; Conservative 7; Mismatches 27; Indels 2; Gaps 2;
```

```
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
DB 649 WYTSSWSECTAECGGGQDRVAVCLN-YDKKVPV-EMCDEAVKPSKQDCNVDDC 701
```

RESULT 6

```
T34395
Hypothetical protein C37C3.6b - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 01-Dec-2000
C:Accession: T34395; T34394
R:Geisel, C.; Bradshaw, H.
Submitted to the EMBL Data Library, July 1996
A:Description: The sequence of C. elegans cosmid C37C3.
A:Reference number: Z21518
A:Accession: T34395
A:Status: preliminary; translated from GB/EMBL/DBJ
```

```
A:Molecule type: DNA
A:Residues: 1-2167 <GEI>
A:Cross-references: EMBL:U64857; PIDN:AAC25868.1; GSPDB:GN00023; CESP:C37C3.6b
A:Experimental source: strain Bristol N2; clone C37C3
A:Accession: T34394
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1555, 'SKF' <GE2>
A:Cross-references: EMBL:U64857; PIDN:AAC25867.1; GSPDB:GN00023; CESP:C37C3.6a
A:Experimental source: strain Bristol N2; clone C37C3
C:Genetics:
A:Gene: CESP:C37C3.6b; CESP:C37C3.6a
A:Map position: 5
A:Introns: 32/3; 104/2; 156/2; 207/1; 459/2; 536/3; 577/2; 1105/3; 1367/1; 1438/1; 155
```

Query Match

```
Best Local Similarity 27.5%; Score 95; DB 2; Length 2167;
Matches 19; Conservative 7; Mismatches 27; Indels 2; Gaps 2;
```

```
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
DB 649 WYTSSWSECTAECGGGQDRVAVCLN-YDKKVPV-EMCDEAVKPSKQDCNVDDC 701
```

RESULT 7

```
T25061
Hypothetical protein T21B6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T25061
```

Ri Cottage, A.
submitted to the EMBL Data Library, November 1995
A;Reference number: Z19975
A;Accession: T25061
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-788 <WIL>
A;Cross-references: EMBL:Z68011; PIDN:CAA92014.1; GSPDB:GN00028; CESP:T21B6.3
A;Experimental source: clone T21B6
C;Genetics:
A;Gene: CESP.T21B6.3
A;Map position: X
A;Introns: 20/1; 47/1; 76/1; 152/1; 735/2; 754/2
Query Match 26.9%; Score 93; DB 2; Length 788;
Best Local Similarity 41.5%; Pred. No. 0.0072;
Matches 22; Conservative 4; Mismatches 15; Indels 12; Gaps 3;
QY 6 WMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHL-----PKPVTVRGCGWAGPC 55
DB 560 WGCSTVCGGVAVRQRTCLG-----GVFGDHLQCGPK-TEQRACDGGGPC 603
RESULT 8
T00017
gene ADAMTS-1 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C;Accession: T00017
R;Kuno, K.; Lizasa, H.; Ohno, S.; Matsushima, K.
J. Biol. Chem. 266, 466-471, 1997
A;Title: The exon/intron organization and chromosomal mapping of the mouse ADAMTS-1 gene
A;Reference number: Z14055; MUID:98110583; PMID:9441751
A;Accession: T00017
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-951 <KUN>
A;Cross-references: EMBL:AB001735; NID:g2809056; PIDN:BAA24501.1; PID:g2809057
A;Experimental source: strain 129SVJ
C;Genetics:
A;Gene: ADAMTS-1
A;Introns: 228/1; 343/3; 388/1; 444/1; 539/3; 602/1; 660/3; 719/2
A;Superfamily: thrombospondin type 1 repeat homology
F;542-598/Domain: thrombospondin type 1 repeat homology <THR3>
Query Match 26.7%; Score 92.5; DB 2; Length 951;
Best Local Similarity 38.2%; Pred. No. 0.0097;
Matches 21; Conservative 3; Mismatches 28; Indels 3; Gaps 1;
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLPKPVTVRGCGWAGPC 55
DB 842 WVIEWGECSKTCGSGWQRRVQCRDINGH---PASECAKEYKPASTRPCADLPC 893
RESULT 9
T47158
hypothetical protein DKFPz762C110.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C;Accession: T47158
R;Blum, H.; Bauersachs, S.; Mewes, H.W.; Weil, B.; Wienann, S.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24379
A;Accession: T47158
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-550 <AAA>
A;Cross-references: EMBL:AL162080
A;Experimental source: adult melanoma (Mewo cell line); clone DKFPz762C110
C;Genetics:
A;Note: DKFPz762C110.1
Query Match 26.4%; Score 91.5; DB 2; Length 550;

Best Local Similarity 40.0%; Pred. No. 0.0077;
Matches 22; Conservative 2; Mismatches 28; Indels 3; Gaps 1;
QY 1 WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCQHLPKPVTVRGCGWAGPC 55
DB 441 WVIEWGECSKTCGSGWQRRVQCRDINGQ---PASECAKEYKPASTRPCADHPC 492
RESULT 10
A42587
thrombospondin 2 precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
C;Accession: A42587; A39851
R;Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.
J. Biol. Chem. 267, 3274-3281, 1992
A;Title: Characterization of mouse thrombospondin 2 sequence and expression during cel.
A;Reference number: A42587; MUID:92147683; PMID:1371115
A;Accession: A42587
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: nucleic acid
A;Residues: 1-1172 <LAH>
A;Cross-references: GB:L07803; GB:M87275; NID:g340421; PIDN:AAA53064.1; PID:g567241
A;Note: sequence extracted from NCBI backbone (NCBI:81502)
R;Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.; Li, P.; Dixit, V.M.
J. Biol. Chem. 266, 12821-12824, 1991
A;Title: A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome.
A;Reference number: A39851; MUID:91302287; PMID:1712771
A;Accession: A39851
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-873 <BOR>
A;Cross-references: GB:M64866; NID:g201994; PIDN:AAA40432.1; PID:g201995
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology;
C;Keywords: calcium binding; glycoprotein
F;319-377/Domain: von Willebrand factor type C repeat homology <VWC>
F;380-431/Domain: thrombospondin type 1 repeat homology <THR1>
F;436-492/Domain: thrombospondin type 1 repeat homology <THR2>
F;493-549/Domain: thrombospondin type 1 repeat homology <THR3>
F;553-588/Domain: thrombospondin type 1 repeat homology <EGF1>
F;652-691/Domain: EGF homology <EGF>
Query Match 25.9%; Score 89.5; DB 2; Length 1172;
Best Local Similarity 35.1%; Pred. No. 0.026;
Matches 20; Conservative 3; Mismatches 11; Indels 23; Gaps 2;
QY 6 WMECSVSCGDIQRR-----RDTCLGPAQAPVPADFCQHLPKPVTVRGCGWAGPC 55
DB 390 WTSCSVTCGSGTQQRGSCDVTSNCLGPSIQ-----TRTCSLGKC 430
RESULT 11
A39804
thrombospondin precursor - chicken
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A39804
R;Lawler, J.; Duquette, M.; Ferro, P.
J. Biol. Chem. 266, 8039-8043, 1991
A;Title: Cloning and sequencing of chicken thrombospondin.
A;Reference number: A39804; MUID:91217026; PMID:2022631
A;Accession: A39804
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1178 <LAW>
A;Cross-references: GB:M60853; NID:g212763; PIDN:AAA51437.1; PID:g212764
C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology <VWC>
F;325-383/Domain: von Willebrand factor type C repeat homology <VWC>
F;386-437/Domain: thrombospondin type 1 repeat homology <THR1>
F;442-498/Domain: thrombospondin type 1 repeat homology <THR2>
F;499-555/Domain: thrombospondin type 1 repeat homology <THR3>
F;658-697/Domain: EGF homology <EGF>

Qy 6 WMECSVSCDGIQRR-----RDTCLGPOAQ 30
Db 390 WTQCSVTCSGSGTQQRGRSCDVTSNTCLGPSIQ 421

Search completed: March 13, 2004, 07:44:39
Job time : 2.02396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:29:54 ; Search time 1.20833 Seconds
(without alignments)

2499.367 Million cell updates/sec

Title: US-09-836-712-2 COPY 1099 1156

Perfect score: 346

Sequence: 1 WHVGTWMECSVSCGDIQRR.....QHLPKPVTTVRGCWAGPC'VGQ 58

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Loss Probability: Minimum Match 0%
Maximum Match 10%

18111 Burrasin

Database : SwissProt 42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			ID	Description
	Score	Match	Length		
1	123.5	35.7	1906	1	AT20 MOUSE
2	121.5	35.1	1081	1	AT18 HUMAN
3	119.5	34.5	1935	1	AT59 HUMAN
4	116.5	33.7	950	1	AT15 HUMAN
5	113.5	32.8	1911	1	AT20 HUMAN
6	110.5	31.9	1211	1	AT52 HUMAN
7	104.5	30.2	1205	1	AT53 HUMAN
8	101.5	29.3	1205	1	AT52 BOVIN
9	100.5	29.0	905	1	AT58 MOUSE
10	99.5	28.8	890	1	AT58 HUMAN
11	98.5	28.5	1593	1	AT12 HUMAN
12	98	28.3	1210	1	AT19 MOUSE
13	96.5	27.9	562	1	AT15 MOUSE
14	95.5	27.6	1213	1	AT52 MOUSE
15	94.5	27.3	1223	1	AT14 HUMAN
16	93.5	27.0	1224	1	AT16 HUMAN
17	93	26.9	1207	1	AT19 HUMAN
18	92.5	26.7	967	1	AT51 RAT
19	92.5	26.7	968	1	AT51 MOUSE
20	92	26.6	1173	1	TSPI XENLA
21	91.5	26.4	967	1	AT51 HUMAN
22	89.5	25.9	1172	1	TSP2 CHICK
23	89	25.7	1178	1	TSP2 MOUSE
24	87	25.1	1095	1	AT17 HUMAN
25	86.5	25.0	1077	1	AT10 HUMAN
26	86.5	25.0	1170	1	TSP2 BOVIN
27	85	24.6	525	1	AT11 HUMAN
28	83	24.0	450	1	AT10 MOUSE
29	82.5	23.8	1172	1	TSP2 HUMAN
30	81.5	23.6	1572	1	BA12 HUMAN
31	81	23.4	1170	1	TSP1 BOVIN
32	81	23.4	1170	1	TSP1 MOUSE
33	80.5	23.3	1093	1	SM58 HUMAN

This SWISS-PROT entry is copyright. It is produced through a collaboration

ALIGNMENTS

```

RESULT 1
AT20_MOUSE STANDARD; PRT; 1906 AA.
ID AT20_MOUSE
AC P59511;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-20 precursor (PC 3.4.24.-) (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-TS 20) (ADAM-
DE TS20).
CS GN ADAMTS20.
CS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_taxid=10090;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1).
RP TISSUE=Fetal brain;
RC MEDLINE=22566039; PubMed=12562771;
RA Ilamazares M., Cal S., Quesada V., Lopez-Otin C.;
RT "Identification and characterization of ADAMTS-20 defines a novel
RT subfamily of metalloproteinases-disintegrins with multiple
RT thrombospondin-1 repeats and a unique GCN domain."
RL J. Biol. Chem. 278:13382-13399(2003).

```

[2]
SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND DISEASE.

STRAIN=DBA/2;
MEDLINE=22806432; PubMed=12925592;
Rao C., Foerzler D., Loftus S.K., Liu S., McPherson J.D.,
Jungers K.A., Apte S., Pavan W.J., Beiler D.R.;
"A defect in a novel ADAMTS family member is the cause of the belted
white-spotting mutation.";
Development 130:4665-4672(2003).
-!- FUNCTION: May play a role in tissue-remodeling process occurring
in both normal and pathological conditions.
-!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
-!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).

- ! - ALTERNATIVE PRODUCTS:

```
Event=Alternative splicing; Named isoforms=2;
```

Name=1; Synonyms=ADAMTS20 B long isoform;

IsId=P59511-1; Sequence=Displayed;
Name=3. 3'-non-transcribed 3' short isoform.

```

Name=2; synonym8=ADAMTS20 A SHORT ISOFORM;
IcdId=PE9511-3; Sequence=VSR 007506 VSR 007507;

```

ISOLID=P393II-2; sequence=VSP_007606; VSP_007607;
-|- TISSUE SPECIFICITY: Expressed at low level in testis and brain.

- PTM: The precursor is cleaved by a furin endopeptidase (By

similarity).

-!- DISEASE: Defects in ADAMTS20 are the cause of the belted (bt)

phenotype. It is a pigmental defe

defect in melanocyte development.

-|- SIMILARITY: Belongs to peptidase family M12B.

- /- SIMILARITY: Contains 1 disintegrin-like domain.
- /- SIMILARITY: Contains 1 CON domain
- /- SIMILARITY: Contains 1 CON domain

```
-|- SIMILARITY: CONTAINS 1 GON domain.
-|- SIMILARITY: CONTAINS 15 TSP type-1 domains.
```

-!- SIMILARITY: CONTAINS 13 OF CYPE-1 DOMAINS

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; AJ512753; CAD54808.3; -
DR EMBL; AV189815; AO74895.1; -
DR EMBL; AV189816; AO74896.1; -
DR MGD; NGI:2660628; Adants20.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Pep_M12B_propep; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 11.
DR SMART; SM00209; TSP1; 14.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00092; TSP1; 13.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT PROPEP 27 249 BY SIMILARITY.
FT CHAIN 250 1906 ADAMTS-20.
FT DOMAIN 250 464 METALLOPROTEASE.
FT DOMAIN 465 552 DISINTEGRIN-LIKE.
FT DOMAIN 553 608 TSP TYPE-1 1.
FT DOMAIN 609 720 CYS-RICH.
FT DOMAIN 721 842 SPACER.
FT DOMAIN 843 901 TSP TYPE-1 2.
FT DOMAIN 906 962 TSP TYPE-1 3.
FT DOMAIN 962 1015 TSP TYPE-1 4.
FT DOMAIN 1017 1074 TSP TYPE-1 5.
FT DOMAIN 1075 1131 TSP TYPE-1 6.
FT DOMAIN 1148 1202 TSP TYPE-1 7.
FT DOMAIN 1203 1260 TSP TYPE-1 8.
FT DOMAIN 1300 1351 TSP TYPE-1 9.
FT DOMAIN 1354 1411 TSP TYPE-1 10.
FT DOMAIN 1412 1465 TSP TYPE-1 11.
FT DOMAIN 1468 1526 TSP TYPE-1 12.
FT DOMAIN 1527 1584 TSP TYPE-1 13.
FT DOMAIN 1585 1648 TSP TYPE-1 14.
FT DOMAIN 1650 1706 TSP TYPE-1 15.
FT DOMAIN 1707 1906 GON.
FT METAL 399 399 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 400 400 BY SIMILARITY.
FT METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 409 409 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 798 798 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 805 805 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1562 1562 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1719 1719 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1759 1759 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1777 1777 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 1424 1425 CS -> VR (in isoform 2).
FT VARSPPLIC 1426 1426 Missing (in isoform 2).
FT CONFLICT 1211 1211 D -> Y (IN REF. 2).
FT CONFLICT 1262 1262 S -> L (IN REF. 2).
FT SEQUENCE 1906 AA; 212041 MW; 0EB2568547B557EA CRC64;

Query Match

35.7%; Score 123.5; DB 1; Length 1906;

Best Local Similarity 40.0%; Pred. No. 2.6e-06;
Matches 22; Conservative 8; Mismatches 24; Indels 1; Gaps 1;
QY 1 MHVGTWMEGVSVCQGIQRDRDTCLGPQAPVADFCQHLPKPVTVRCGWAGPC 55
DB 1416 WTRGPKWCSASCCKGVKREVLCI-DQFQKLEKYCSHLHKPETHKACRSGR 1469
RESULT 2
AT18_HUMAN STANDARD; PRT; 1081 AA.
ID AT18_HUMAN
AC Q8TE60;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-18 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 18) (ADAM-TS 18) (ADAM-TS18).
GN ADAMTS18.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT Cloning, expression analysis, and structural characterization of
RT seven novel human ADAMTS, a family of metalloproteinases with
RT disintegrin and thrombospondin-1 domains.";
RL Gene 283:49-62(2002).
CC -|- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -|- TISSUE SPECIFICITY: Expressed in fetal lung, liver, and kidney
CC and in adult brain, prostate, submaxillary gland, and endothelium.
CC -|- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -|- SIMILARITY: Belongs to peptidase family M12B.
CC -|- SIMILARITY: Contains 1 disintegrin-like domain.
CC -|- SIMILARITY: Contains 4 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

KW

Repeat; Extracellular matrix.

FT SIGNAL 1 47
FT PROPEP 48 284
FT CHAIN 285 1081
FT DOMAIN 285 497
FT DOMAIN 498 577
FT DOMAIN 588 643
FT DOMAIN 644 749
FT DOMAIN 750 876
FT DOMAIN 877 931
FT DOMAIN 933 991
FT DOMAIN 993 1048
FT SITE 254 254
FT METAL 435 435
FT ACT SITE 436 436
FT METAL 439 439
FT METAL 445 445
FT CARBOHYD 151 151
FT CARBOHYD 190 190
FT CARBOHYD 313 313
FT CARBOHYD 744 744
FT CARBOHYD 837 837
FT CARBOHYD 908 908
SQ SEQUENCE 1081 AA; 119655 MW; 0438BF65676461E CRC64;

Query Match 35.1%; Score 121.5; DB 1; Length 1081;
Best Local Similarity 38.2%; Pred. No. 2.6e-06;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;

Qy 1 WHVGTWNECSVSGDGIQRDRDTCGLGQAAQVPADFCQHLKPVTVGCGWAGPC 55
Db 994 WSLGPNQSCSKTCGRGVKRELLCKGAAEF-LPESQCTSLPRPELQSGCVLGR 1047

RESULT 3
AT9S HUMAN STANDARD; PRT: 1935 AA.
AC Q9P2N4; Q9NR29;
DT 16-OCT-2001 (Rel. 40, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-9 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 9) (ADAM-TS 9) (ADAM-TS9).
GN ADAMTS9 OR KIA1312.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 3).
RC TISSUE=Fetal;
RX MEDLINE=20396138; PubMed=10936055;
RA Clark M.E., Kelner G.S., Turbeville L.A., Boyer A., Arden K.A.,
RA Maki R.A.;
RT "ADAMTS 9, a novel member of the ADAM-TS/Metalloprotein gene
RT family.";
RL Genomics 67:343-350(2000).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND FUNCTION.
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wight T.N., Leduc R., Apte S.S.;
RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RT subfamily related to Ctenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513(2003).
RN [3]
RP SEQUENCE OF 159-1935 FROM N.A. (ISOFORM 2).
RC TISSUE=Brain;
RX MEDLINE=20181126; PubMed=10718198;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:65-73(2000).

CC -!- FUNCTION: Cleaves the large aggregating proteoglycans, aggrecan
CC and versican.
CC -!- CATALYTIC ACTIVITY: Cleaves aggrecan at the 1838-Glu-Ala-1839
CC site and versican at the 1428-Glu-Ala-1429 site.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=ADAMTS-9B;
CC IsoId=Q9P2N4-3; Sequence=Displayed;
CC Name=2; Synonyms=Long;
CC IsoId=Q9P2N4-1; Sequence=VSP_007548, VSP_007549;
CC Note=May result from the retention of an intron in the cDNA
CC leading to a premature stop codon;
CC Name=3; Synonyms=Short;
CC IsoId=Q9P2N4-2; Sequence=VSP_005499, VSP_005500;
CC -!- TISSUE SPECIFICITY: Highly expressed in all fetal tissues.
CC Expressed in a number of adult tissues with highest expression in
CC heart, placenta and skeletal muscle.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC -!- DOMAIN: The ancillary domains, including the TSPs domain, are
CC required for specific extracellular localization and for its
CC versicanase and aggrecanase activities.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF261918; AAF89106.1; -;
CC EMBL: AF488803; AAO15765.1; -;
CC EMBL: AB037733; BAA92550.1; -;
CC HSP: P15167; IATL.
CC MEROPS: M12.021; -;
CC Genes: HGNC:13202; ADAMTS9.
CC MIM: 605421; -;
CC GO: GO:0008237; F:metallopeptidase activity; TAS.
CC GO: GO:0007275; P:development; TAS.
CC GO: GO:0006516; P:glycoprotein catabolism; TAS.
CC InterPro: IPR001762; Disintegrin.
CC InterPro: IPR001818; Pept_M10A_M12B.
CC InterPro: IPR006025; Pept_M2n_BS.
CC InterPro: IPR001590; Peptidase_M12B.
CC InterPro: IPR002870; Peptidase_M12B_N.
CC InterPro: IPR000884; TSP1.
CC Pfam: PF01562; Pept_M12B_proprep; 1.
CC Pfam: PF01421; Reprolysin; 1.
CC SMART: SM00090; tsp_1; 11.
CC SMART: SM00209; TSP1; 12.
CC PROSITE: PS00215; ADAM_MEPPO; 1.
CC PROSITE: PS00545; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE: PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE: PS00214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE: PS00092; TSP1; 14.
CC PROSITE: PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Repeat; Extracellular matrix; Alternative splicing.
CC SIGNAL 1 28
CC PROPEP 19 287
CC CHAIN 288 1935
CC DOMAIN 293 499
CC METALLOPROTEASE.

SEQUENCE FROM N.A.
MEDLINE=21856482; PubMed=11867212;
Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
Lopez-Otin C.;
"Cloning, expression analysis, and structural characterization of
seven novel human ADAMTSs, a family of metalloproteinases with
disintegrin and thrombospondin-1 domains.";
Gene 283:49-62(2002).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed in fetal liver and kidney, but not
in any of the adult tissues examined.
CC -!- DOMAIN: The spacer domain and the TSR type-1 domains are important
for a tight interaction with the extracellular matrix (By
similarity).
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
or send an email to license@isb-sib.ch).

CC EMBL; AJ315733; CAC86014.1; -;
CC MEROPS; M12.025; -;
CC Genew; HGNC:18305; ADAMTS15.
CC MIM; 607509; -;
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 3.
CC SMART; SM03209; TSP1; 3.
CC PROSITE; PS00215; ADAM_MPRO; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS0092; TSP1; 3.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC KW Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 212 BY SIMILARITY.
FT CHAIN 213 950 ADAMTS-15.
FT DOMAIN 213 427 METALLOPROTEASE.
FT DOMAIN 428 515 DISINTEGRIN-LIKE.
FT DOMAIN 516 571 TSP TYPE-1 1.
FT DOMAIN 572 700 CYS-RICH.
FT DOMAIN 701 838 SPACER.
FT DOMAIN 839 895 TSP TYPE-1 2.
FT DOMAIN 896 949 TSP TYPE-1 3.
FT SITE 174 174 CYSTEINE SWITCH (POTENTIAL).
FT METAL 361 361 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 362 362 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 365 365 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 371 371 N-LINKED (GLNAC...) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLNAC...) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLNAC...) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLNAC...) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLNAC...) (POTENTIAL).
FT CARBOHYD 950 950 AA; 103286 MW; 5DFEB18285CCCC3B CRC64;
SQ

Query Match 33.7%; Score 116.5; DB 1; Length 950;
Best Local Similarity 41.8%; Pred. No. 8.6e-06;
Matches 23; Conservative 5; Mismatches 24; Indels 3; Gaps 2;

Oy 1 WHVGTWMECSVCGDGIORRRDTCGLPQAPVPADFCQHLKPVTVRGWAGPC 55
Db 843 WVAGSWGPCASCGSLQKRAVDCRSAGQRTVPA--CDAHRPVEIQAC-GEPC 894

RESULT 5
AT20_HUMAN STANDARD; PRT; 1911 AA.
AC P59510;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-20 precursor [EC 3.4.24.-] (A disintegrin and
DE metalloproteinase with thrombospondin motifs 20) (ADAM-
DE TS20).
GN ADAMTS20.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=22513925; PubMed=12514189;
RA Somerville R.P., Longpre J.-M., Jungers K.A., Engle J.M., Ross M.,
RA Evanko S., Wright T.N., Leduc R., Apce S.S.;
RT "Characterization of ADAMTS-9 and ADAMTS-20 as a distinct ADAMTS
RT subfamily related to Caenorhabditis elegans GON-1.";
RL J. Biol. Chem. 278:9503-9513(2003).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Liver;
RX MEDLINE=22566039; PubMed=12562771;
RA Llanazares M., Cal S., Ouesada V., Lopez-Otin C.;
RT "Identification and characterization of ADAMTS-20 defines a novel
RT subfamily of metalloproteinases-disintegrins with multiple
RT thrombospondin-1 repeats and a unique GON domain.";
RL J. Biol. Chem. 278:13382-13389(2003).
CC -!- FUNCTION: May play a role in tissue-remodeling process occurring
CC in both normal and pathological conditions.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59510-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59510-2; Sequence=VSP_007106, VSP_007107, VSP_007108;
CC -!- TISSUE SPECIFICITY: Very sparingly expressed, although is detected
CC at low levels in testis, prostate, ovary, heart, placenta, lung
CC and pancreas. Overexpressed in several brain, colon and breast
CC carcinomas.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 GON domain.
CC -!- SIMILARITY: Contains 15 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF488804; AA015766.1; -.
CC EMBL; AJ515153; CAD56159.3; -.

DR DR EMBL; AJ515154; CAD56160.2; -.
DR Genew; HGNC:17178; ADAMTS20
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Rep_M12B_propep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp_1; 11.
DR SMART; SM00209; TSP1; 12.
DR PROSITE; PS0215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS0092; TSP1; 12.
DR PROSITE; PS0142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Alternative splicing.
FT SIGNAL 1 21 POTENTIAL.
FT PROPEP 22 253 BY SIMILARITY.
FT CHAIN 254 1911 ADAMTS-20.
FT DOMAIN 254 467 METALLOPROTEASE.
FT DOMAIN 468 555 DISINTEGRIN-LIKE.
FT DOMAIN 556 611 TSP TYPE-1 1.
FT DOMAIN 612 723 CYS-RICH.
FT DOMAIN 724 846 SPACER.
FT DOMAIN 847 905 TSP TYPE-1 2.
FT DOMAIN 906 962 TSP TYPE-1 3.
FT DOMAIN 967 1024 TSP TYPE-1 4.
FT DOMAIN 1025 1074 TSP TYPE-1 5.
FT DOMAIN 1077 1136 TSP TYPE-1 6.
FT DOMAIN 1153 1207 TSP TYPE-1 7.
FT DOMAIN 1208 1265 TSP TYPE-1 8.
FT DOMAIN 1305 1357 TSP TYPE-1 9.
FT DOMAIN 1359 1417 TSP TYPE-1 10.
FT DOMAIN 1418 1476 TSP TYPE-1 11.
FT DOMAIN 1477 1532 TSP TYPE-1 12.
FT DOMAIN 1536 1589 TSP TYPE-1 13.
FT DOMAIN 1590 1653 TSP TYPE-1 14.
FT DOMAIN 1655 1711 TSP TYPE-1 15.
FT DOMAIN 1712 1911 GON.
FT METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT 404 404 BY SIMILARITY.
FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 92 92 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 702 702 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 809 809 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 870 870 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1062 1062 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1457 1457 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1543 1543 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1573 1573 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1764 1764 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1782 1782 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1853 1853 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 683 693 THDICVQGCQM -> SYNIDNCVCLK (in isoform
FT 2).
FT VARSPLIC 1430 1504 /FTId=VSP_007106.
FT CSACGGRKRVFCIDQFORKLEDTNCSQVQKPTHKAC
FT RSVRCPKSWKNECVTCGSGVQQRDVYCRK -> BDL
FT KVKLPORITILWELMKNIFCHGKSHMYLINVTDHLLYP
FT RHCDPEITETFLSLWSLQFTWGLDKYKNSL (in
FT isoform 2).
FT /FTId=VSP_007107.
FT Missing (in isoform 2).
FT T -> V (IN REF. 2).
FT E -> Y (IN REF. 2).

FT CONFLICT 371 371 T -> M (IN REF. 2).

FT CONFLICT 456 456 D -> E (IN REF. 2).

FT CONFLICT 648 648 I -> V (IN REF. 2).

FT CONFLICT 752 752 MISSING (IN REF. 2).

FT CONFLICT 818 821 LILQ -> ILIE (IN REF. 2).

FT CONFLICT 908 913 GMLAK -> WHVIG (IN REF. 2).

FT CONFLICT 1316 1316 S -> Q (IN REF. 2).

FT CONFLICT 1328 1328 T -> R (IN REF. 2).

FT CONFLICT 1817 1817 T -> A (IN REF. 2; CAD56159).

FT CONFLICT 1882 1882 E -> Q (IN REF. 2; CAD56159).

SQ SEQUENCE 1911 AA; 214656 MW; CF592B22D032B250 CRC64;

Query Match 32.8%; Score 113.5; DB 1; Length 1911;

Best Local Similarity 35.7%; Pred. No. 3.7e-05;

Matches 20; Conservative 9; Mismatches 26; Indels 1; Gaps 1;

QY 1 WHVGTWECVSGDGIQRERDCLGFAQAPVAFDCQHLPKVTVRGWAGPCV 56

DB 1478 WKANSWNECVSGGVQ-ORDVYCRLLKGVGVVEEMCDQSTRCSQRRCSQCV 1532

RESULT 6

ATIS2_HUMAN STANDARD; PRT; 1211 AA.

ID ATIS2_HUMAN

AC O95450;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase with thrombospondin motifs 2) (ADAM-15) (Procollagen I/II

DE amino-propeptidase processing enzyme) (Procollagen I N-proteinase) (PC

DE I-NP) (Procollagen N-endopeptidase) (pNPI).

GN ADAMTS2 OR PCINP OR PCFNI.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORMS LPNFI AND SPNPI), AND DISEASE.

RX TISSUE=Skin;

RC MEDLINE=99347935; PubMed=10417273;

RA Colige A., Sieron A.L., Li S.-W., Schwarze U., Petty E.,

RA Wartecki W., Wilcox W., Krakow D., Cohn D.H., Reardon W.,

RA Byers P.H., Lapiere C.M., Prockop D.J., Nussgens B.V.;

RT "Human Ehlers-Danlos syndrome type VII C and bovine dermatosparaxis are caused by mutations in the procollagen I N-proteinase gene."

RL Am. J. Hum. Genet. 65:308-317(1999).

CC -1- FUNCTION: Cleaves the propeptides of type I and II collagen prior to fibril assembly. Does not act on type III collagen. May also play a role in development that is independent of its role in collagen biosynthesis.

CC -1- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain alpha-1(I) at Pro-|-Gln and of alpha-1(II) and alpha-2(I) chains at Ala-|-Gln.

CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).

CC -1- SUBUNIT: May belong to a multimeric complex. Binds specifically to collagen type XIV (by similarity).

CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).

CC -1- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Name=LPNFI;

CC IsoId=O95450-1; Sequence=Displayed;

CC Name=SPNPI;

CC IsoId=O95450-2; Sequence=VSP_005497, VSP_005498;

CC Note=Has no significant N-procollagen peptidase activity; tendon

CC -1- TISSUE SPECIFICITY: Expressed at high level in skin, bone, tendon and aorta and at low levels in thymus and brain.

CC -1- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.

CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).

CC -1- DISEASE: Defects in ADAMTS2 are the cause of Ehlers-Danlos

CC syndrome type VIIC (EDS VIIC) [MIM:225410]. EDS VIIC is a recessively inherited connective-tissue disorder characterized clinically by severe skin fragility and joint hypermobility and biochemically by the presence in skin of procollagen incompletely processed at the N-terminus.

CC -1- SIMILARITY: Belongs to peptidase family M12B.

CC -1- SIMILARITY: Contains 1 disintegrin-like domain.

CC -1- SIMILARITY: Contains 1 PLAC domain.

CC -1- SIMILARITY: Contains 4 TSP type-1 domains.

CC -1- CAUTION: Has sometimes been referred to as ADAMTS3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AJ003125; CAA05880.1; --

DR MEROPS; M12.301; --

DR Genew; HGNC:218; ADAMTS2.

DR MIM; 604539; --

DR MIM; 225410; --

DR GO; GO:0008133; F:collagenase activity; TAS.

DR GO; GO:0008237; F:metallopeptidase activity; TAS.

DR InterPro; IPR001762; Disintegrin.

DR InterPro; IPR006025; Pept_M_Zn_BS.

DR InterPro; IPR001590; Peptidase_M12B.

DR InterPro; IPR002870; Peptidase_M12B_N.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP1.

DR Pfam; PF01562; Pep_M12B_propep; 1.

DR Pfam; PF01421; Reprolysin; 1.

DR Pfam; PF00900; tsp_1; 4.

DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00209; TSP1; 4.

DR PROSITE; PS02115; ADAM_MEPPO; 1.

DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.

DR PROSITE; PS02114; DISINTEGRIN_2; FALSE_NEG.

DR PROSITE; PS0900; PLAC; 1.

DR PROSITE; PS0092; TSP1; 4.

DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.

DR PROSITE; PS00142; ZINC_PROTEASE; Zinc; Signal; Glycoprotein; Zymogen; KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen; KW Repeat; Collagen degradation; Extracellular matrix; KW Alternative splicing; Ehlers-Danlos syndrome.

FT SIGNAL 1 29 POTENTIAL.

FT PROPEP 30 253 BY SIMILARITY.

FT CHAIN 254 1211 ADAMTS-2.

FT DOMAIN 254 479 METALLOPROTEASE.

FT DOMAIN 480 560 DISINTEGRIN-LIKE.

FT DOMAIN 561 616 TSP TYPE-1 1.

FT DOMAIN 618 722 CYS-RICH.

FT SITE 691 693 CELL ATTACHMENT SITE (POTENTIAL).

FT DOMAIN 723 851 SPACER.

FT DOMAIN 854 912 TSP TYPE-1 2.

FT DOMAIN 914 971 TSP TYPE-1 3.

FT DOMAIN 975 1029 TSP TYPE-1 4.

FT DOMAIN 1059 1097 PLAC.

FT DOMAIN 40 43 POLI-ALA.

FT DOMAIN 185 188 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 408 408 BY SIMILARITY.

FT ACT SITE 409 409 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 412 412 ZINC (CATALYTIC) (BY SIMILARITY).

FT METAL 418 418 ZINC (CATALYTIC) (BY SIMILARITY).

FT CARBOHYD 112 112 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 949 949 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 993 993 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1031 1031 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1098 1098 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1145 1145 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1150 1150 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT VARSPLIC 544 566 HCFKGHCWILTPDILKRDGWSGA -> FRPGAVAHACYPST
FT LGQOGSWIA (in isoform SpNP1).
FT /FTID=VSP_005497.
FT Missing (in isoform SpNP1).
FT /FTID=VSP_005498.
SQ SEQUENCE 1211 AA; 134722 MW; BCEBEF25C23CAD2D CRC64;

Query Match 31.9%; Score 110.5; DB 1; Length 1211;
Best Local Similarity 40.3%; Pred. No. 5.3e-05;
Matches 25; Conservative 4; Mismatches 14; Indels 19; Gaps 4;

Qy 1 WHVGTWMECSVSCGDGIQRDRDTCGLGQQAQVPE---AD----FCQHLKPVTVRGCMAG 53
Db 979 WRAGPWSQSVTCGNGTQER-----PVCRTADDSFGICQB-ERPETARTCLRG 1026

Qy 54 PC 55
Db 1027 PC 1028

RESULT 7
ID AT3S HUMAN STANDARD; PRT; 1205 AA.
AC O15072; O9BXZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-3 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 3) (ADAM-TS 3) (Procollagen II
DE amino-propeptide processing enzyme) (Procollagen II N-proteinase) (PC
DE II-NP).
GN ADAMTS3 OR KIAA0366.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE OF 1-227 FROM N.A.
RX MEDLINE=21402912; PubMed=11408482;
RA Fernandes R.J., Hirohata S., Engle J.M., Colige A., Cohn D.H.,
RA Eyre D.R., Apte S.S.;
RT "Procollagen II amino propeptide processing by ADAMTS-3. Insights on
RT dermatosparaxis.";
RL J. Biol. Chem. 276:31502-31509 (2001).
RN [2]
RP SEQUENCE OF 5-1205 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150 (1997).
CC -!- FUNCTION: Cleaves the propeptides of type II collagen prior to
CC fibril assembly. Does not act on types I and III collagens.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Found in cartilage and skin.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- CAUTION: Has sometimes been referred to as ADAMTS4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF247668; AAK38400.1; -;
DR EMBL; AB002364; BAA20821.1; -;
DR MEROPS; M12.220; -;
DR Genew; HGNC:219; ADAMTS3.
DR MIM; 605011; -;
DR GO; GO:0005578; C:extracellular matrix; NAS.
DR GO; GO:0004222; F:metalloendopeptidase activity; NAS.
DR GO; GO:0008270; F:zinc ion binding; NAS.
DR GO; GO:0030574; P:collagen catabolism; NAS.
DR GO; GO:0030199; P:collagen fibril organization; NAS.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP 1.
DR Pfam; PF01562; Pep_M12B_Dropep; 1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsp 1; 4.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS00215; ADAM_MPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00300; PLAC; 1.
DR PROSITE; PS00092; TSP1; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL; 1 20
FT PROPEP 21 249
FT CHAIN 250 1205
FT DOMAIN 250 469
FT DOMAIN 470 550
FT DOMAIN 551 606
FT DOMAIN 608 712
FT DOMAIN 713 844
FT DOMAIN 845 905
FT DOMAIN 906 965
FT DOMAIN 966 1014
FT DOMAIN 1015 1054
FT DOMAIN 245 249
FT METAL 398 398
FT ACT_SITE 399 399
FT METAL 402 402
FT METAL 408 408
FT CARBOHYD 83 83
FT CARBOHYD 119 119
FT CARBOHYD 242 242
FT CARBOHYD 345 345
FT CARBOHYD 475 475
FT CARBOHYD 814 814
FT CARBOHYD 942 942
SQ SEQUENCE 1205 AA; 135570 MW; EB07B286FC85PB87 CRC64;

Query Match 30.2%; Score 104.5; DB 1; Length 1205;
Best Local Similarity 38.2%; Pred. No. 0.00026;
Matches 21; Conservative 5; Mismatches 20; Indels 9; Gaps 2;

Qy 1 WHVGTWMECSVSCGDGIQRDRDTCGLGQQAQVPEADFCQHLKPVTVRGCMAGPC 55
Db 970 WKTGPMSECVTCGEGTEVRQVLGRA-----GDHCDG-EKPSVRACQLPPC 1015

RESULT 8
ID AT3S BOVIN STANDARD; PRT; 1205 AA.
AC P79331;

16-OCT-2001 (Rel. 40, Created)
 16-OCT-2001 (Rel. 40, Last sequence update)
 15-MAR-2004 (Rel. 43, Last annotation update)
 ADAMTS-2 precursor (EC 3.4.24.14) (A disintegrin and metalloproteinase with thrombospondin motifs 2) (ADAM-TS2) (Procollagen I/II amino-propeptide processing enzyme) (Procollagen I N-proteinase) (PC I-NP) (Procollagen N-endopeptidase) (5NP1).
 ADAMTS2 OR NPI.
 Bos taurus (Bovine).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Artiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Skin;
 MEDLINE=97225960; PubMed=9122202;
 Colige A., Li S.W., Sieron A.L., Nusgens B.V., Prockop D.J., Lapiere C.M.;
 "cDNA cloning and expression of bovine procollagen I N-proteinase: a new member of the superfamily of zinc-metalloproteinases with binding sites for cells and other matrix components.";
 Proc. Natl. Acad. Sci. U.S.A. 94:2374-2379 (1997).
 [2]
 PARTIAL SEQUENCE.
 MEDLINE=95348096; PubMed=7622483;
 Colige A., Beschin A., Sany B., Goebels Y., Van Beeumen J., Nusgens B.V., Lapiere C.M.;
 "Characterization and partial amino acid sequencing of a 107-kDa procollagen I N-proteinase purified by affinity chromatography on immobilized type XIV collagen.";
 J. Biol. Chem. 270:16724-16730 (1995).
 -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior to fibril assembly. Does not act on type III collagen. May also play a role in development that is independent of its role in collagen biosynthesis.
 -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain at alpha-1(I) at Pro-Gln and of alpha-1(II) and alpha-2(I) chains at Ala-Gln.
 -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
 -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to collagen type XIV.
 -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular matrix (By similarity).
 -!- TISSUE SPECIFICITY: Enzymatic activity is detected at high level in all type I collagen-rich tissues such as skin, bones, tendons and aorta and at low level in brain and thymus. The mRNA levels were disproportionately high in heart, liver, retina and muscle.
 -!- DOMAIN: The spacer domain and the TSP type-1 domains are important for a tight interaction with the extracellular matrix.
 -!- PTM: The N-terminus is blocked.
 -!- PTM: The precursor is cleaved by a furin endopeptidase (By similarity).
 -!- DISEASE: Defects in ADAMTS2 are the cause of dermatosparaxis, a recessively inherited disorder characterized by severe skin fragility and biochemically by the presence in skin of procollagen incompletely processed at the N-terminus.
 -!- SIMILARITY: Belongs to peptidase family M12B.
 -!- SIMILARITY: Contains 1 disintegrin-like domain.
 -!- SIMILARITY: Contains 1 PLAC domain.
 -!- SIMILARITY: Contains 4 TSP type-1 domains.
 -!- CAUTION: Has sometimes been referred to as ADAMTS3.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; X96389; CAA65253.1; -.
 PIR; T18517; T18517.

DR MEROPS; M12.301; -.
 DR InterPro; IPR001762; Disintegrin.
 DR InterPro; IPR006025; Pept_M_Zn_BS.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR002870; Peptidase_M12B_N.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF01562; Pep_M12B_propep; 1.
 DR Pfam; PF01421; Reprolysin; 1.
 DR Pfam; PF00090; TSP_1; 4.
 DR SMART; SM00209; TSP1; 4.
 DR PROSITE; PS00215; ADAM_MPRO; 1.
 DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
 DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
 DR PROSITE; PS00900; PLAC; 1.
 DR PROSITE; PS00092; TSP1; 4.
 DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
 KW Hydrolase; Metalloproteinase; Zinc; Signal; Glycoprotein; Zymogen;
 Repeat; Collagen degradation; Extracellular matrix.
 FT SIGNAL 1 28
 FT PROPEP 29 253
 FT CHAIN 254 1205
 FT DOMAIN 254 473
 FT DOMAIN 474 554
 FT DOMAIN 555 610
 FT DOMAIN 612 716
 FT DOMAIN 717 845
 FT DOMAIN 848 906
 FT DOMAIN 908 968
 FT DOMAIN 969 1023
 FT DOMAIN 1053 1091
 FT DOMAIN 31 35
 FT DOMAIN 177 180
 FT METAL 402 402
 FT ACT_SITE 403 403
 FT METAL 406 406
 FT SITE 412 412
 FT SITE 685 687
 FT CARBOHYD 104 104
 FT CARBOHYD 245 245
 FT CARBOHYD 942 942
 FT CARBOHYD 943 943
 FT CARBOHYD 987 987
 FT CARBOHYD 1025 1025
 FT CARBOHYD 1092 1092
 FT CARBOHYD 1139 1139
 FT CARBOHYD 1144 1144
 SQ SEQUENCE 1205 AA; 133887 MW; 7B5B232A45320371 CRC64;
 Query Match 29.3%; Score 101.5; DB 1; Length 1205;
 Best Local Similarity 38.2%; Pred. No. 0.00056;
 Matches 21; Conservative 6; Mismatches 23; Indels 5; Gaps 2;
 QY 1 WHVGTWMECSVSGDG-QRRRDTCIGPQAQAPVPADFQHLPKPVTVRGCMWAGPC 55
 Db 973 WRAGWSQCSVTGNGTQERPVLC-----RTADDSFGVCRE-ERPETATCRILGPC 1022
 RESULT 9
 ID_ATS8_MOUSE STANDARD; PRT; 905 AA.
 AC PS7110;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2).
 GN ADAMTS8.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

```

RX MEDLINE=20079168; PubMed=10610729;
RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
RL mouse chromosome 9 and human chromosome 11.";
CC -!- FUNCTION: Has anti-angiogenic properties (By similarity).
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Expressed specifically in adult lung and heart
CC and low expression during mouse development.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF175282; AAF25805.1; -.
CC HSP; P34179; I1AG.
CC MEROPS; M12.226; -.
CC MGD; MGI:1353468; Adamts8.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC Pfam; PF01421; Reprolysin; 1.
CC Pfam; PF00090; tsp_1; 2.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00209; TSP1; 2.
CC PROSITE; PS00215; ADAM_MPRO; 1.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS0092; TSP1; 2.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL; 1 28 POTENTIAL.
FT PROPEP; 29 228 BY SIMILARITY.
FT CHAIN; 229 905 ADAMTS-8.
FT DOMAIN; 229 452 METALLOPROTEASE.
FT DOMAIN; 453 541 DISINTEGRIN-LIKE.
FT DOMAIN; 542 597 TSP TYPE-1 1.
FT DOMAIN; 599 705 CYS-RICH.
FT DOMAIN; 706 847 SPACER.
FT DOMAIN; 848 904 TSP TYPE-1 2.
FT METAL; 378 378 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE; 379 379 BY SIMILARITY.
FT METAL; 382 379 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL; 388 388 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD; 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD; 480 480 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD; 506 506 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD; 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 905 AA; 98879 MW; 124D4132B33A0CAE CRC64;

Query Match
Best Local Similarity 29.0%; Score 100.5; DB 1; Length 905;
Matches 20; Conservative 5; Mismatches 27; Indels 3; Gaps 1;

QY 1 WHVGTWMECVSGDGIQRREDTCLGPOAQVPADFCQHLKPKVTVRCWAGPC 55
Db 852 WVLGDWSECFPCRGSMQRRVTECRDPGQA--SDTCDALKPEDAKPCGSGPC 903
```

```

RESULT 10
RTS8_HUMAN STANDARD; PRT; 890 AA.
ID ATIS8_HUMAN
AC Q9UP79; Q9NZS0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 8) (ADAM-TS 8) (ADAM-TS8) (METH-2)
DE (METH-8).
GN ADAMTS8 OR METH2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=lung;
RX MEDLINE=99367466; PubMed=10438512;
RA Vazquez F., Hastings G., Ortega M.-A., Lane T.F., Oikemus S.,
RA Lombardo M., Iruela-Arispe M.L.;
RT "METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new
RT family of proteins with angio-inhibitory activity.";
RL J. Biol. Chem. 274:23349-23357(1999).
RN [2]
RP SEQUENCE OF 195-440 FROM N.A.
RX MEDLINE=20079168; PubMed=10610729;
RA Georgiadis K.E., Hirohata S., Seldin M.F., Apte S.S.;
RT "ADAM-TS8, a novel metalloprotease of the ADAM-TS family located on
RT mouse chromosome 9 and human chromosome 11.";
RL Genomics 62:312-315(1999).
CC -!- FUNCTION: Has anti-angiogenic properties.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in adult and fetal lung, lower
CC expression in brain, placenta, heart, stomach and fetal brain and
CC kidney.
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 2 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF060153; AAD48081.1; -.
CC EMBL; AF175283; AAF25806.1; -.
CC HSP; P34179; I1AG.
CC MEROPS; M12.226; -.
CC Genew; HGNC:224; ADAMTS8.
CC MIM; 605175; -.
CC GO; GO:0005178; F.integrin binding; TAS.
CC GO; GO:0008237; F.metalloproteinase activity; TAS.
CC GO; GO:0007345; P.embryogenesis and morphogenesis; TAS.
CC GO; GO:0008285; P.negative regulation of cell proliferation; TAS.
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
```

DR PFAM: PF01421; Reprolysin; 1.
DR PFAM: PF00090; TSP1; 2.
DR PRINTS: PR01705; TSP1REPEAT.
DR SMART: SM00209; TSP1; 2.
DR PROSITE: PS00215; ADAM MEPRO; 1.
DR PROSITE: PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE: PS00214; DISINTEGRIN_2; FALSE NEG.
DR PROSITE: PS00092; TSP1; 2.
DR PROSITE: PS00142; ZINC_PROTEASE; 1.
KW Hydrolyase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix; Heparin-binding.
FT SIGNAL 1 27
FT PROPEP 28 214
FT CHAIN 215 890
FT DOMAIN 215 430
FT DOMAIN 439 526
FT DOMAIN 527 582
FT DOMAIN 584 690
FT DOMAIN 691 832
FT DOMAIN 834 889
FT DOMAIN 202 205
FT METAL 364 364
FT ACT SITE 365 365
FT METAL 368 368
FT METAL 374 374
FT CARBOHYD 345 345
FT CARBOHYD 401 401
FT CARBOHYD 466 466
FT CARBOHYD 491 491
FT CARBOHYD 600 600
FT CONFLICT 195 195
FT CONFLICT 413 440
SQ SEQUENCE 890 AA; 96671 MW; 57D70EE03D5739D3 CRC64;

Query Match 28.8; Score 99.5; DB 1; Length 890;
Best Local Similarity 39.2; Pred. No. 0.00071;
Matches 21; Conservative 4; Mismatches 27; Indels 3; Gaps 1;

QY 1 WHVGTMECSVSGDGHORRDTCLGPOAQPVPADFCQHLKPKVTYVRCWGAGPC 55
DB 837 WVLGDWSECSSTCGAGWQRRTVECRDFSGQASAT---CNKALKPDKAPCESQLC 888

RESULT 11
AT12_HUMAN STANDARD; PRT; 1593 AA.
AC P58397;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-12 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
DE with thrombospondin motifs 12) (ADAM-TS 12) (ADAM-TS12).
GN ADAMTS12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal lung;
RC MEDLINE=21264577; PubMed=11279086;
RX Cal S., Argueelles J.M., Fernandez P.L., Lopez-Otin C.;
RA "Identification, characterization, and intracellular processing of
RT ADAM-TS12, a novel human disintegrin with a complex structural
RT organization involving multiple thrombospondin-1 repeats.";
RL J. Biol. Chem. 276:17932-17940 (2001).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
CC matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed exclusively in fetal lung. Is widely
CC expressed in gastric carcinomas and in cancer cells of diverse
CC origin.

FT CARBOHYD 685 685 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 790 790 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1300 1300 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1320 1320 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1371 1371 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1378 1378 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1503 1503 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1593 AA; 177545 MW; 0799F48E63BD83A3 CRC64;

Query Match 28.5%; Score 98.5; DB 1; Length 1593;
Best Local Similarity 24.1%; Pred. No. 0.0016;
Matches 27; Conservative 5; Mismatches 17; Indels 63; Gaps 3;

QY 1 WHVGTWECVSCDGIQRRDTCI-----QAPVPADFQCHLPKPVTVRG 50
DB 827 WQYGHWTCTGTTGIRTAHCIRKGRGMVATFDPETQNGRQKCKEAKCPPRW 886

QY 26 -----GPOA-----QAPVPADFQCHLPKPVTVRG 50
DB 887 AGEWEACATCGPRGKKTCTVLCQTMVSDQALPPTD-CQHLAKPKILLSC 937

RESULT 12
AT19_MOUSE
ID AT19_MOUSE STANDARD; PRT; 1210 AA.
AC P59509;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-19 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 19) (ADAM-TS19) (ADAM-TS19).
GN ADAMTS19.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TIGSUE=Ovary;
RX MEDLINE=22505168; PubMed=12617826;
RA Menke D.B., Page D.C.;
RT "Sexually dimorphic gene expression in the developing mouse gonad."
RL Gene Expr. Patterns 2:359-367(2002).
CC -1- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed predominantly in fetal ovary, low
levels of expression is also detected in kidney, heart, skeletal
muscle, lung and testis.
CC -1- DEVELOPMENTAL STAGE: Expression is strongest in anterior and
ventral regions of the ovary at 12.5 and 13.5 dpc before becoming
more uniform.
CC -1- PMW: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -1- SIMILARITY: Belongs to peptidase family M12B.
CC -1- SIMILARITY: Contains 1 disintegrin-like domain.
CC -1- SIMILARITY: Contains 5 TSP type-1 domains.
CC -1- SIMILARITY: Contains 1 PLAC domain.
CC -1- CAUTION: By homology with the human sequence, it is uncertain
whether Met-1 or Met-5 is the initiator.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AY135183; AAN10155.1; --

DR MGD; MGI:2442875; Adamts19.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR001818; Rept M10A_M12B.
DR InterPro; IPR006025; Rept M Zn BS.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSPI.
DR Pfam; PF01562; Pep M12B_propep; 1.
DR Pfam; PF01421; Reprylisin; 1.
DR Pfam; PF00090; tsp.1; 5.
DR SMART; SM00508; ACE; 1.
DR SMART; SM00209; TSPI; 5.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00546; CYSTEINE SWITCH; FALSE NEG.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE NEG.
DR PROSITE; PS00214; DISINTEGRIN_2; FALSE NEG.
DR PROSITE; PS00900; PLAC; 1.
DR PROSITE; PS00092; TSPI; 5.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR K0 Hydrolyase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Extracellular matrix.
FT SIGNAL 1 30 POTENTIAL.
FT PROPEP 31 319 BY SIMILARITY.
FT CHAIN 320 1210 ADAMTS-19.
FT DOMAIN 320 548 METALLOPROTEASE.
FT DOMAIN 549 636 DISINTEGRIN-LIKE.
FT DOMAIN 637 689 TSP TYPE-1 1.
FT DOMAIN 690 793 CYS-RICH.
FT DOMAIN 794 917 SPACER.
FT DOMAIN 918 978 TSP TYPE-1 2.
FT DOMAIN 979 1040 TSP TYPE-1 3.
FT DOMAIN 1042 1086 TSP TYPE-1 4.
FT DOMAIN 1090 1147 TSP TYPE-1 5.
FT DOMAIN 1163 1202 PLAC.
FT SITE 297 297 CYSTEINE SWITCH (POTENTIAL).
FT METAL 485 485 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT SITE 486 486 BY SIMILARITY.
FT METAL 489 489 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 495 495 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 800 800 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 910 910 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 931 931 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1012 1012 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1210 AA; 134560 MW; 0AB812ABAB4BB7A2 CRC64;

Query Match 28.3%; Score 98; DB 1; Length 1210;
Best Local Similarity 34.5%; Pred. No. 0.0014;
Matches 20; Conservative 7; Mismatches 25; Indels 6; Gaps 2;

QY 1 WHVGTWECVSCDGIQRRDTCIQLGPOAQA-----PVPADFQCHLPKPVTVR-GCWA 52
DB 1043 WEAGWSECVKCGKGVHRVTRCTNPKKCVLSTPRREAECDSDISKCVWVGWDS 1100

RESULT 13
AT15_MOUSE
ID AT15_MOUSE STANDARD; PRT; 562 AA.
AC P59384;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE ADAMTS-15 (EC 3.4.24.-) (A disintegrin and metalloproteinase with
thrombospondin motifs 15) (ADAM-TS15) (ADAM-TS15) (Fragment).
GN ADAMTS15.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.C., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton B., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J.J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalrus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Cleaves the propeptides of type I and II collagen prior
to fibril assembly. Does not act on type III collagen. May also
play a role in development that is independent of its role in
collagen biosynthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: Cleaves the N-propeptide of collagen chain
alpha-1(I) at Pro1-Gln and of alpha-1(II) and alpha-2(I) chains
at Ala1-Gln.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- SUBUNIT: May belong to a multimeric complex. Binds specifically to
collagen type XIV (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -!- DOMAIN: The spacer domain and the TSP type-1 domains are important
for a tight interaction with the extracellular matrix.
CC -!- PTM: The precursor is cleaved by a furin endopeptidase (By
similarity).
CC -!- SIMILARITY: Belongs to peptidase family M12B.
CC -!- SIMILARITY: Contains 1 disintegrin-like domain.
CC -!- SIMILARITY: Contains 1 PLAC domain.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AK040370; BAC30572.1; -;
DR EMBL; BC046456; AAH46456.1; -;
DR MGD; MGI:1347355; Adams2.
DR InterPro; IPR001762; Disintegrin.
DR InterPro; IPR006025; Pept_M12B_N.
DR InterPro; IPR001590; Peptidase_M12B.
DR InterPro; IPR002870; Peptidase_M12B_N.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF01562; Pep_M12B_prop; 1.
DR Pfam; PF01421; Reprolysin; 1.
DR Pfam; PF00090; tsp.1; 4.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
DR PROSITE; PS00214; DISINTEGRIN_2; FALSE_NEG.
DR PROSITE; PS00900; PLAC; 1.
DR PROSITE; PS00092; TSP1; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
KW Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
KW Repeat; Collagen degradation; Extracellular matrix.
FT SIGNAL 1 28
FT PROPEP 29 260
FT CHAIN 29 1213
FT DOMAIN 261 471
FT DOMAIN 480 560
FT DOMAIN 561 617
FT DOMAIN 618 722
FT DOMAIN 723 851
FT DOMAIN 855 913
FT DOMAIN 915 975
FT DOMAIN 976 1030

FT DOMAIN 1060 1098
FT SITE 692
FT METAL 409
FT ACT SITE 410
FT METAL 413
FT METAL 419
FT CARBOHYD 111
FT CARBOHYD 252
FT CARBOHYD 949
FT CARBOHYD 950
FT CARBOHYD 994
FT CARBOHYD 1032
FT CARBOHYD 1099
FT CARBOHYD 1147
FT CARBOHYD 1152
SQ SEQUENCE 1213 AA; 135298 MW; B27431E00443EDB5 CRC64;
Query Match 27.6%; Score 95.5; DB 1; Length 1213;
Best Local Similarity 36.4%; Pred. No. 0.0028;
Matches 20; Conservative 6; Mismatches 24; Indels 5; Gaps 2;
QY 1 MWVGTWECVSCGGDIQRRTCLGPOQAQYPADFCQHLKPEVTVRGCGWAGPC 55
Db 980 WRAGSWSQCVTCGNGTQRPVLCRTADDNFGV----CRE-ERPETARICRLAPC 1029
RESULT 15
AT14 HUMAN STANDARD; PRT; 1223 AA.
ID AT14 HUMAN STANDARD; PRT; 1223 AA.
AC Q8WXS8; Q8TE55; Q8TEY8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE ADAMTS-14 precursor (BC 3.4.24.-) (A disintegrin and metalloproteinase
with thrombospondin motifs 14) (ADAM-TS 14) (ADAM-TS14).
GN ADAMTS14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM A).
RX MEDLINE=21638061; PubMed=11779638;
RA Bolz H., Ramirez A., von Bredelow B., Kubisch C.;
RT "Characterization of ADAMTS14, a novel member of the ADAMTS
metalloproteinase family";
RL Biochim. Biophys. Acta 1522:221-225(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM A).
RC TISSUE=Fetal lung;
RX MEDLINE=21856482; PubMed=11867212;
RA Cal S., Obaya A.J., Llamazares M., Garabaya C., Quesada V.,
RA Lopez-Otin C.;
RT "Cloning, expression analysis, and structural characterization of
seven novel human ADAMTSs, a family of metalloproteinases with
disintegrin and thrombospondin-1 domains";
RL Gene 283:49-62(2002).
RN [3]
RP SEQUENCE OF 29-1223 FROM N.A. (ISOFORMS B; C AND D), AND
RP ALTERNATIVE PROMOTER USAGE.
RX MEDLINE=21839041; PubMed=11741898;
RA Collige A., Vandenberghe I., Thiry M., Lambert C.A., Van Beeumen J.,
RA Li S.-W., Prockop D.J., Lapiere C.M., Nusgens B.V.;
RT "Cloning and characterization of ADAMTS-14, a novel ADAMTS displaying
high homology with ADAMTS-2 and ADAMTS-3";
RL J. Biol. Chem. 277:5756-5766(2002).
CC -!- FUNCTION: Has a aminoprocollagen type I activity processing
activity in the absence of ADAMTS2. Seems to be synthesized as a
latent enzyme that requires activation to display aminoprocollagen
peptidase activity.
CC -!- SUBCELLULAR LOCATION: Secreted. Associated with the extracellular
matrix (By similarity).
CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative promoter;
CC Comment=2 isoforms, A (shown here) and B, are produced by use of
CC alternative promoters;
CC Event=Alternative splicing; Named isoforms=4;
CC Name=A;
CC IsoId=Q8WXS8-1; Sequence=Displayed;
CC Name=B;
CC IsoId=Q8WXS8-2; Sequence=VSP_006958;
CC Name=C;
CC IsoId=Q8WXS8-3; Sequence=VSP_006958, VSP_005501;
CC Notes=Produced by alternative splicing of isoform B;
CC Name=D;
CC IsoId=Q8WXS8-4; Sequence=VSP_005501;
CC Notes=Produced by alternative splicing of isoform A;
CC TISSUE SPECIFICITY: Expressed in retina and at low levels in
CC brain, lung and placenta. High expression in fetal tissues.
CC !- DOMAIN: The spacer domain and the TSP type-1 domains are important
CC for a tight interaction with the extracellular matrix (By
CC similarity).
CC !- PTM: The precursor is cleaved by a furin endopeptidase (By
CC similarity).
CC !- SIMILARITY: Belongs to peptidase family M12B.
CC !- SIMILARITY: Contains 1 disintegrin-like domain.
CC !- SIMILARITY: Contains 1 PLAC domain.
CC !- SIMILARITY: Contains 4 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF358666; AAL40229.1; -;
CC EMBL; AJ345098; CAC87943.1; -;
CC EMBL; AF366351; AAL79814.1; -;
CC MEROPS; M12.024; -;
CC Genew; HGNC:14899; ADAMTS14.
CC MIM; 607506; -;
CC InterPro; IPR001762; Disintegrin.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR001590; Peptidase_M12B.
CC InterPro; IPR002870; Peptidase_M12B_N.
CC InterPro; IPR008884; TSP1.
CC InterPro; IPR008085; TSP1.
CC Pfam; PF01562; Pep_M12B_propep; 1.
CC Pfam; PF01421; Repolysin; 1.
CC Pfam; PF00090; tsp.1; 4.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SMO0209; TSP1; 4.
CC PROSITE; PS0215; ADAM_MEPRO; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; FALSE_NEG.
CC PROSITE; PS00427; DISINTEGRIN_1; FALSE_NEG.
CC PROSITE; PS0214; DISINTEGRIN_2; FALSE_NEG.
CC PROSITE; PS0300; PLAC; 1.
CC PROSITE; PS00092; TSP1; 4.
CC PROSITE; PS00142; ZINC_PROTEASE; FALSE_NEG.
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;
CC Collagen degradation; Repeat; Extracellular matrix;
CC Alternative splicing; Alternative promoter usage.
CC SIGNAL 1 22 POTENTIAL.
CC FT SIGNAL 1 22 BY SIMILARITY.
CC FT PROPEP 23 252
CC FT CHAIN 253 1223 ADAMTS-14.
CC FT DOMAIN 253 460 METALLOPROTEASE.
CC FT DOMAIN 461 551 DISINTEGRIN-LIKE.
CC FT DOMAIN 552 607 TSP TYPE-1 1.
CC FT DOMAIN 608 729 CYS-RICH.
CC FT DOMAIN 730 846 SPACER.
CC FT DOMAIN 847 907 TSP TYPE-1 2.
CC FT DOMAIN 908 967 TSP TYPE-1 3.
CC FT DOMAIN 968 1022 TSP TYPE-1 4.

FT DOMAIN 1059 1097
FT DOMAIN 875 878
FT DOMAIN 1100 1221
FT ACT_SITE 339
FT METAL 338
FT METAL 402 402
FT METAL 408 408
FT CARBOHYD 109 109
FT CARBOHYD 475 475
FT CARBOHYD 941 941
FT CARBOHYD 1027 1027
FT VARSPLIC 1 68
FT VARSPLIC 368
FT CONFLICT 868 868
FT CONFLICT 884 884
FT CONFLICT 901 901
FT CONFLICT 923 923
FT CONFLICT 1024 1024
FT SEQUENCE 1223 AA; 133871 MW; 36394AC4D92F170F CRC64;
Query Match 27.3%; Score 94.5; DB 1; Length 1223;
Best Local Similarity 33.3%; Pred. No. 0.0036;
Matches 21; Conservative 9; Mismatches 16; Indels 17; Gaps 3;
QY 1 WHVGTWEGSVSCDGIQRRTTC-----LGPQAQAPVPADFCQHLKPKVTVRGCWAGP 54
DB 972 WRIGAMQCSATCGEGIQQRVVCRTNANSLG-----HCEG-DRPDTVQVCSLPA 1020
QY 55 CVG 57
DB 1021 CGG 1023

Search completed: March 13, 2004, 07:39:57
Job time : 1.20833 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:30:14 ; Search time 5.73958 Seconds
(without alignments)
3188.394 Million cell updates/sec

Title: US-09-836-712-2_COPY_1099_1156
Perfect score: 346
Sequence: 1 WHVGTWMECSVSGDGIQRR.....QHLKPVTIVRGWAGPCVQ 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 25:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rviro:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	346	100.0	1427	Q96L37	Q96L37 homo sapien
2	125.5	36.3	640	11 Q8BSM2	Q8BSM2 mus musculus
3	122.5	35.4	237	4 Q9HSE6	Q9HSE6 homo sapien
4	122.5	35.4	424	4 Q8N643	Q8N643 homo sapien
5	122.5	35.4	701	11 Q8CDB8	Q8CDB8 mus musculus
6	118	34.1	761	6 Q9SLQ2	Q9SLQ2 macaca fasc
7	118	34.1	898	4 Q9URF24	Q9URF24 homo sapien
8	117.5	34.0	1092	11 Q8BKA1	Q8BKA1 mus musculus
9	116.5	33.7	235	11 Q91YP4	Q91YP4 mus musculus
10	116.5	33.7	1036	11 Q80T21	Q80T21 mus musculus
11	116	33.5	1280	11 Q9EPX2	Q9EPX2 mus musculus
12	115	33.2	1023	4 Q9ULI17	Q9ULI17 homo sapien
13	114	32.9	766	4 P82987	P82987 homo sapien
14	114	32.9	1235	4 Q95428	Q95428 homo sapien
15	113.5	32.8	417	4 Q9BXK3	Q9BXK3 homo sapien
16	112	32.4	2165	5 Q19791	Q19791 caenorhabdi

17	109.5	31.6	454	4 Q96H81	Q96H81 homo sapien
18	109.5	31.6	538	4 Q9H8E4	Q9H8E4 homo sapien
19	109.5	31.6	622	4 Q9H8X0	Q9H8X0 homo sapien
20	109.5	31.6	658	11 Q8BLE5	Q8BLE5 mus musculus
21	107.5	31.1	432	4 Q9NPM2	Q9NPM2 homo sapien
22	107	30.9	4123	4 Q75851	Q75851 homo sapien
23	104.5	30.2	364	4 Q96AY5	Q96AY5 homo sapien
24	102.5	29.6	1572	5 Q44938	Q44938 haemochus
25	101.5	29.3	759	6 Q8H2M8	Q8H2M8 equus caball
26	101.5	29.3	1043	5 P90884	P90884 caenorhabdi
27	100.5	29.0	183	11 Q99JF1	Q99JF1 mus musculus
28	98.5	28.5	2174	5 Q9GQR0	Q9GQR0 drosophila
29	98.5	28.5	2772	5 Q9VAV4	Q9VAV4 drosophila
30	98.5	28.5	2776	5 Q869A0	Q869A0 drosophila
31	98.5	28.5	2898	5 Q86829	Q86829 drosophila
32	98	28.3	3198	5 Q9U8G8	Q9U8G8 manduca sex
33	96.5	27.9	340	11 Q91J256	Q91J256 mus musculus
34	96.5	27.9	1009	11 Q8BKX1	Q8BKX1 mus musculus
35	96.5	27.9	1600	11 Q811B3	Q811B3 mus musculus
36	96	27.7	117	11 Q8BNI0	Q8BNI0 mus musculus
37	95	27.5	685	6 Q9TTS5	Q9TTS5 bos taurus
38	95	27.5	1487	5 Q8MPV5	Q8MPV5 caenorhabdi
39	95	27.5	1558	5 Q817I0	Q817I0 caenorhabdi
40	95	27.5	2167	5 Q76840	Q76840 caenorhabdi
41	95	27.5	5146	6 Q8SPW4	Q8SPW4 bos taurus
42	94	27.2	403	5 Q8MRW9	Q8MRW9 drosophila
43	94	27.2	1353	5 Q9V9Q7	Q9V9Q7 drosophila
44	93	26.9	821	5 Q22631	Q22631 caenorhabdi
45	93	26.9	975	5 Q7Z291	Q7Z291 caenorhabdi

ALIGNMENTS

RESULT 1

Q96L37 PRELIMINARY; PRT; 1427 AA.
ID Q96L37
AC Q96L37
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Von Willebrand factor-cleaving protease precursor.
GN ADAMTS13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX PubMed=11557746;
RA Zheng X., Chung D., Takayama T.K., Majerus E.M., Sadler J.E., Fujikawa K.;
RA "Structure of von Willebrand Factor-cleaving Protease (ADAMTS13), a Metalloprotease Involved in Thrombotic Thrombocytopenic Purpura.";
RL J. Biol. Chem. 276:41059-41063 (2001).
DR MEROPS; M12.241; -
DR GO; HGNC:1366; ADAMTS13.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001590; Peptidase M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR000884; TSF1.
DR Pfam; PF01421; Repolysin; 1.
DR Pfam; PF00090; tsep1; 4.
DR SMART; SM00209; TSF1; 7.
DR PROSITE; PS50215; ADAM_MEPRO; 1.
DR PROSITE; PS50092; TSF1; 4.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Protease; Signal.
FT SIGNAL 1 33 POTENTIAL.

FT CHAIN 75 1427 VON WILLEBRAND FACTOR-CLEAVING PROTEASE.
SQ SEQUENCE 1427 AA; 153632 MW; EB1BC3AABCI1A4442 CRC64;

Query Match 100.0%; Score 346; DB 4; Length 1427;
Best Local Similarity 100.0%; Pred. No. 4e-35;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRDTCCLGPOAQPVPADFCQHLPKPVTVRGCGWAGPCVQ 59
DB 1076 WHVGTWMECSVSCGDIQRRDTCCLGPOAQPVPADFCQHLPKPVTVRGCGWAGPCVQ 1133

RESULT 2
QBSM2 PRELIMINARY; PRT; 640 AA.
AC QBSM2;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Forelimb;
RA MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
DR EMBL; AK031246; BAC27317.1; -.
DR PIR; PT0566; PT0566.
DR PIR; PT0633; PT0633.
DR InterPro; IPR003598; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig C2.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00047; Ig; 3.
DR Pfam; PF00090; tsp.1; 2.
DR SMART; SM00409; IG; 3.
DR SMART; SM00408; IGC2; 3.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS50835; IG-Like; 3.
DR PROSITE; PS50092; TSP1; 2.
KW Hypothetical protein.
SQ SEQUENCE 640 AA; 69371 MW; 0729FC3372DB900A CRC64;

Query Match 36.3%; Score 125.5; DB 11; Length 640;
Best Local Similarity 38.2%; Pred. No. 1.8e-07;
Matches 29; Conservative 6; Mismatches 22; Indels 19; Gaps 4;

QY 1 WHVGTWMECSVSCGDIQRRDTCCLGPOA---CAPVPADFCQHLPKPVTVRGCGWAGPCVQ 50
DB 427 WMVTSWASCTSCGGVQTRRVTCOKLKASGISTPVSNWCSQLAKRPVDTQACNQLCV 486

QY 51 -WA-----GPCVG 57
DB 487 EWAFSSWQCNGPCIG 502

RESULT 3
Q9HBS6 PRELIMINARY; PRT; 237 AA.
AC Q9HBS6;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gu J.R., Wan D.F., Zhao X.T., Zhou X.M., Jiang H.Q., Zhang P.P.,
RA Qin W.X., Huang Y., Qiu X.K., Qian L.F., He L.P., Li H.N., Yu Y.,
RA Yu J., Han L.H.;
RA "Novel Human CDNA clones with function of inhibiting cancer cell
RT growth";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF217974; AAG17217.1; -.
DR Genew; HGNC:19706; TSP1.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 2.
DR SMART; SM00209; TSP1; 3.
DR PROSITE; PS50092; TSP1; 3.
KW Hypothetical protein.
SQ SEQUENCE 237 AA; 25659 MW; FB73ED571107026C CRC64;

Query Match 35.4%; Score 122.5; DB 4; Length 237;
Best Local Similarity 41.0%; Pred. No. 1.6e-07;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQPVPADFCQHLPKPVTVRGCGWAGPCVQ 57
DB 77 WYTGPMGECSECGSGTQRDIICVSKLGTGFNVTSFSPN-CSHLPRPPALQ-----PCQG 130

QY 58 Q 58
DB 131 Q 131

RESULT 4
Q8N643 PRELIMINARY; PRT; 424 AA.
AC Q8N643;
DT 01-OCT-2002 (T-EMBLrel. 22, Created)
DT 01-OCT-2002 (T-EMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027478; AAH27478.1; -.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp.1; 4.
DR SMART; SM00209; TSP1; 6.
DR PROSITE; PS50092; TSP1; 5.
SQ SEQUENCE 424 AA; 45774 MW; C4C51297C98C75EB CRC64;

Query Match 35.4%; Score 122.5; DB 4; Length 424;
Best Local Similarity 41.0%; Pred. No. 2.8e-07;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQPVPADFCQHLPKPVTVRGCGWAGPCVQ 57
DB 264 WYTGPMGECSECGSGTQRDIICVSKLGTGFNVTSFSPN-CSHLPRPPALQ-----PCQG 317

QY 58 Q 58
DB 318 Q 318

RESULT 5
Q8CDB8 PRELIMINARY; PRT; 701 AA.
AC Q8CDB8;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN 8430403M15RIK OR AW743315.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK030699; BAC27087.1; -
 DR MGD; MGI:2141719; 8430403M15RIK.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 8.
 DR SMART; SM00209; TSP1; 8.
 DR PROSITE; PS50092; TSP1; 8.
 KW Hypothetical protein.
 SQ SEQUENCE 701 AA; 77924 MW; C854E5D0704805F CRC64;

 Query Match 35.4%; Score 122.5; DB 11; Length 701;
 Best Local Similarity 43.6%; Pred. No. 4.6e-07;
 Matches 24; Conservative 5; Mismatches 25; Indels 1; Gaps 1;

 QY 1 WHVGTWMECSVSCGDGQRRDTCGLGPOAQAPVPADFCQHLKPKVTVRGCGWAGPC 55
 DB 238 WSTGPMSSCSVSCGRGHKRNVCYLAQDG-SHLESDNCKHLPKHGRRRCRGRC 291

 RESULT 6
 Q95LQ2 PRELIMINARY; PRT; 761 AA.
 AC Q95LQ2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
 OC Cercopitheciinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 RA Terao K., Sugano S.,
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
 RT libraries."
 RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AB072736; BAB69705.1; -
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003598; IG_c2.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00047; ig; 3.
 DR Pfam; PF00090; tsp_1; 2.
 DR SMART; SMC0408; IG2; 1.
 DR SMART; SM00209; TSP1; 3.
 DR PROSITE; PS50835; IG LIKE; 3.
 DR PROSITE; PS50092; TSP1; 3.
 KW Hypothetical protein; Immunoglobulin domain.
 SQ SEQUENCE 761 AA; 84196 MW; CEB3F558FCB8AF06 CRC64;

 Query Match 34.1%; Score 118; DB 6; Length 761;
 Best Local Similarity 36.4%; Pred. No. 1.9e-06;
 Matches 20; Conservative 8; Mismatches 27; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDGQRRDTCGLGPOAQAPVPADFCQHLKPKVTVRGCGWAGPC 55
 DB 688 WHAGPMKPCTAACGRGLQSRKVDCHTRSCKPVAEHRHCVCQKKKPVSWRHCLGPSC 742

 RESULT 7
 Q9UFZ4 PRELIMINARY; PRT; 898 AA.
 ID Q9UFZ4;
 AC Q9UFZ4;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein (Fragment).
 GN DKFZP434H204.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AL110226; CAB53684.1; -
 DR PIR; T14764; T14764.
 DR InterPro; IPR000884; TSP1.
 DR Pfam; PF00090; tsp_1; 5.
 DR SMART; SM00209; TSP1; 7.
 DR PROSITE; PS50092; TSP1; 6.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 898 AA; 96731 MW; 38532F1E56F0A11B CRC64;

 Query Match 34.1%; Score 118; DB 4; Length 898;
 Best Local Similarity 42.9%; Pred. No. 2.2e-06;
 Matches 24; Conservative 7; Mismatches 23; Indels 2; Gaps 2;

QY 1 WHVGTWMECSVSCGDGQRRDTCGLGPOAQAPVPADFCQHLKPKVTVRGCGWAGPC 55
 DB 37 WHYGPWTKCTVTCGRGVQRNVYCLERQA-GFVDEEHCDPLGRPDQQRKCSQPC 91

 RESULT 8
 Q8BKAL PRELIMINARY; PRT; 1092 AA.
 ID Q8BKAL;
 AC Q8BKAL;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical metalloprotease.
 GN E130314N14RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Eye;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573 (2002).
 DR EMBL; AK533851; BAC35556.1; -
 DR MGD; MGI:2442600; E130314N14RIK.
 DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001590; Peptidase_M12B.
 DR InterPro; IPR000884; TSP1.
 DR InterPro; IPR008085; TSP1.
 DR Pfam; PF01421; Reprolysin; 1.

RA Buchner D.A., Meisler M.H.;
RA "tsc1", a widely expressed gene containing seven thrombospondin type I
PT MEDLINE=262393009; PubMed=12706863;
RA

DR	EMBL; AF237652; AAKI5041.1; -.
DR	InterPro; IPR000884; TSPI.
DR	Pfam; PF00090; tsp_1; 6.
DR	SMART; SMO0209; TSP1; 7.
DR	PROSITE; PS50092; TSP1; 5.
KW	Repeat; Signal
FT	SIGNAL 1 26
FF	CHAIN 27 766
FT	DOMAIN 79 123
FT	TSP TYPE-1 1.
FT	DOMAIN 422 474
FT	TSP TYPE-1 2.
FT	DOMAIN 482 528
FT	TSP TYPE-1 3.
FT	DOMAIN 568 625
FT	TSP TYPE-1 4.
FT	DOMAIN 648 703
FT	TSP TYPE-1 5.
FT	DOMAIN 707 759
FT	TSP TYPE-1 6.
FT	NON TER 766
FSQ	SEQUENCE 766 AA; 85987 MW; A42613E87AE91719 CRC64;
Query Match 32.9%; Score 114; DB 4; Length 766;	
Best Local Similarity 44.0%; Pred.No. 6.le-06;	
Matches 22; Conservative 7; Mismatches 19; Indels 2;	
QY	1 WHVGTWMECVSCGDGIQRREDTCLGPAQAQVPADFCQHLKPVTVRGC 50 : : : : : : : : : :
DB	707 WHVGSWGECSATCGVGIGIQTDVVCLHP-GETPAPPSECRD-EKPHALQAC 754 : : : : : : : : :
RESULT 14	
O95428	PRELIMINARY; PRT; 1235 AA.
ID O95428 AC O95428	
DT 01-WAY-1999 (TrEMBLrel. 10, Created)	
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)	
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE Hypothetical protein.	
DE Homo sapiens (Human).	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX NCBI_TaxID=9606;	
XP [1]	
RP SEQUENCE FROM N.A.	
RA Rowen L., Madan A., Qin S., Abbasi N., Dors M., Ratcliffe A.,	
RA Madan A., Dickhoff R., Shafer T., James R., Lusky S., Hood L.;	
RL "Complete sequence of the gene for presenilin 1.";	
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	
R EMBL; AF109907; AAC97963.1; -.	
DR HSP; P12111; 2KNT.	
DR GO; GO:0004867; F-serine protease inhibitor activity; IEA.	
DR InterPro; IPR007110; IG-like.	
DR InterPro; IPR003598; IG_c2.	
DR InterPro; IPR002223; Kunitz_BPTI.	
DR InterPro; IPR000884; TSP1.	
DR Pfam; PF00047; ig_3.	
DR Pfam; PF00014; Kunitz_BPTI; 1.	
DR Pfam; PF00090; tsp_1; 5.	
DR PRINTS; PR00759; BASICPTASE.	
DR ProDom; PD00222; Kunitz_BPTI; 1.	
DR SMART; SMO0408; Igc2; 3.	
DR SMART; SMO0131; KU; 1.	
DR SMART; SMO0209; TSP1; 5.	
DR PROSITE; PS00280; BPTI_KUNITZ_1; 1.	
DR PROSITE; PS0279; BPTI_KUNITZ_2; 1.	
DR PROSITE; PS50835; IG LIKE; 3.	
DR PROSITE; PS50092; TSP1; 5.	
KW Hypothetical protein; Immunoglobulin domain; Protease inhibitor;	
KW Serine protease inhibitor.	
SQ SEQUENCE 1235 AA; 133477 MW; A0B44CCE4F3BE350 CRC64;	
Query Match 32.9%; Score 114; DB 4; Length 1235;	
Best Local Similarity 46.4%; Pred.No. 9.9e-06;	
Matches 26; Conservative 3; Mismatches 21; Indels 6;	
QY	1 WHVGTWMECVSCGDGIQRREDTCLGPAQAQVPADFCQHLKPVTVRGCWAGP : : : : : : : : :

DB 510 WHVGTWGLCSKSCSGTRRRQVICAIGP-----PSHCGSLQH-SKPDVPEPCNTQPC 560

RESULT 15

Q9BXY3 PRELIMINARY; PRT; 417 AA.
AC Q9BXY3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Thrombospondin.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Mao Y., Xie Y., Zhou Z., Zhao S., Wang W., Huang Y., Wang S.,
RA Tang R., Chen X., Wu C.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF251058; AAK34948.1; -;
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00090; tsp_1; 3.
DR SMART; SM00209; TSP1; 5.
DR PROSITE; PS50092; TSP1; 4.
SQ SEQUENCE 417 AA; 46274 MW; EF82550656555DC9 CRC64;
Query Match 32.8%; Score 113.5; DB 4; Length 417;
Best Local Similarity 42.1%; Pred.No. 3.8e-06;
Matches 24; Conservative 8; Mismatches 20; Indels 5; Gaps 2;
QY 6 WMECSVSCGDIQERRDTC---LGPQAQAPVPADFCQHLKPVTIVRCGWAGPCVQ 58
DB 265 WSACTVTCGVGTQVIRVRCQVLLSFOSVADLPIDECG-PKPASQACVAGPCSGE 320

Search completed: March 13, 2004, 07:43:20
Job time : 6.73958 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:28:59 ; Search time 8.48854 Seconds
(without alignments)
1930.574 Million call updates/sec

Title: US-09-836-712-2_COPY_1099_1156

Perfect score: 346

Sequence: 1 WHVGTWMECSVSCDGIQRR.....QLLPKPTVTRGCWAGPCVQG 58

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1980s:*
- 3: Geneseq2000s:*
- 4: Geneseq2000s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346	100.0	1297	AAO16618	Human von
2	346	100.0	1323	AAO16620	Human von
3	346	100.0	1353	AAE24449	Human von
4	346	100.0	1353	AAO16617	Human von
5	346	100.0	1378	AAO16619	Human von
6	346	100.0	1416	ABO04153	Human ADA
7	346	100.0	1427	AAE24450	Human von
8	346	100.0	1427	ADD94038	Human aggr
9	346	100.0	1445	ABO08125	Human PMW
10	122.5	35.4	237	ABO06011	Human can
11	122.5	35.4	358	ADG09952	Novel pro
12	122.5	35.4	372	AAE03698	Human gen
13	122.5	35.4	372	ADC22076	Human pro
14	122.5	35.4	491	ABO09336	ADAM prot
15	122.5	35.4	491	ABU60626	Human nov
16	122.5	35.4	513	ABU54545	Human NOV
17	122.5	35.4	645	ABO10454	Human cDN
18	122.5	35.4	645	ABP67041	Human pol
19	122.5	35.4	767	ADE08579	Novel pro
20	122.5	35.4	1074	AAO30837	Human cel
21	122.5	35.4	1088	AAO30829	Human cel
22	122.5	35.1	283	ABO82162	Human NOV
23	121.5	35.1	367	AAE03699	Human gen
24	121.5	35.1	367	ADC22077	Human pro
25	121.5	35.1	369	AAE09696	Human gen

ALIGNMENTS

RESULT 1

AAO16618

ID AAO16618 standard; protein; 1297 AA.

XX

AC AAO16618;

XX

DT 15-MAY-2003 (first entry)

XX

DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #5.

XX

KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;

KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;

KW myocardial infarction; cerebral infarction; arteriosclerosis;

KW platelet thrombosis; stenosis.

XX

OS Homo sapiens.

XX

XX WO200288366-A1.

XX

PD 07-NOV-2002.

XX

PF 25-APR-2002; 2002WO-JP004141.

XX

PR 25-APR-2001; 2001JP-00128342.

PR

PR 27-JUL-2001; 2001JP-00227510.

PR

PR 28-SEP-2001; 2001JP-00302977.

XX

XX 25-JAN-2002; 2002JP-00017596.

XX

PA (KAGA) CHEMO-SERO-THERAPEUTIC RES INST.

XX

XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;

XX WPI; 2003-120479/11.

XX N-PSDB; ABT32584.

DR

DR von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and

XX PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for

PT developing drugs for e.g. myocardial infarction and cerebral infarction.

XX

XX Claim 4; Page 92-101; 144pp; Japanese.

PS

XX The invention comprises the amino acid and coding sequence of a von

CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of

CC the invention are useful in the diagnosis and treatment of

CC thrombocytopenic purpura, and in developing drugs for myocardial

CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,

CC and stenosis. The present amino acid sequence represents a human von

CC Willebrand factor (vWF)-cleaving enzyme-related protein

```
XX
SQ Sequence 1297 AA;
Query Match 100.0%; Score 346; DB 6; Length 1297;
Best Local Similarity 100.0%; Pred. No. 4.1e-32;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 58
Db 1002 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 1059

RESULT 2
ID AAO16620 standard; protein; 1323 AA.
XX
AC AAO16620;
XX
XX
DT 15-MAY-2003 (first entry)
XX
DE Human von Willebrand factor (vWF)-cleaving enzyme-related protein #7.
XX
XX Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
KW myocardial infarction; cerebral infarction; arteriosclerosis;
KW platelet thrombosis; stenosis.
XX
OS Homo sapiens.
XX
XX WO200288366-A1.
XX
XX
PD 07-NOV-2002.
XX
XX
PF 25-APR-2002; 2002WO-JP004141.
XX
XX
PR 25-APR-2001; 2001JP-00128342.
XX
PR 27-JUL-2001; 2001JP-00227510.
XX
PR 28-SEP-2001; 2001JP-00302977.
XX
PR 25-JAN-2002; 2002JP-00017596.
XX
XX
PA (KAGA ) CHEMO-SERO-THERAPEUTIC RES INST.
XX
XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
PI WPI; 2003-120479/11.
XX
DR N-PSDB; ABT32586.
XX
XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
PT supplementary therapy for, thrombotic thrombocytopenic purpura, and for
PT developing drugs for e.g. myocardial infarction and cerebral infarction.
XX
XX
PS Claim 4; Page 112-121; 144pp; Japanese.
XX
CC The invention comprises the amino acid and coding sequence of a von
CC Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
CC the invention are useful in the diagnosis and treatment of
CC thrombocytopenic purpura, and in developing drugs for myocardial
CC infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
CC and stenosis. The present amino acid sequence represents a human von
CC Willebrand factor (vWF)-cleaving enzyme-related protein
XX
SQ Sequence 1323 AA;
Query Match 100.0%; Score 346; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 4.1e-32;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 58
Db 1027 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 1084

RESULT 3
```

```
AAE24449
ID AAE24449 standard; protein; 1353 AA.
XX
AC AAE24449;
XX
XX
DT 04-OCT-2002 (first entry)
XX
DE Human von Willebrand factor-cleaving protease fragment #2.
XX
XX Human; von Willebrand factor-cleaving protease; vWF-cp; therapy; enzyme;
KW transgenic animal; immunisation; thromboembolic disease; preeclampsia;
KW thrombotic thrombocytopenic purpura; TTP; Henoch-Schönlein purpura;
KW thrombosis; neonatal thrombocytopenia; haemolytic-uraemic syndrome;
KW transgenic; anticoagulant.
XX
XX Homo sapiens.
XX
XX WO200242441-A2.
XX
XX
PD 30-MAY-2002.
XX
XX
PF 20-NOV-2001; 2001WO-EP013391.
XX
XX
PR 22-NOV-2000; 2000US-00721254.
XX
PR 12-APR-2001; 2001US-00833328.
XX
XX (BAXT ) BAXTER AG.
XX
XX Laemmle B, Gerritsen HE, Furlan M, Turecek P, Schwarz H;
XX Scheiflinger F, Antoine G, Kerschbaumer R, Tagliavacca L;
XX Zimmermann K, Voelkel D;
XX WPI; 2002-479950/51.
XX
XX Novel isolated or substantially purified von Willebrand factor-cleaving
XX protease, useful for producing preparation for therapy of thrombosis and
XX thromboembolic disease such as thrombotic thrombocytopenic purpura.
XX
XX Claim 1; Page 64-68; 93pp; English.
XX
XX The invention relates to an isolated or substantially pure von Willebrand
XX factor-cleaving protease (vWF-cp) polypeptide. vWF-cp is useful for
XX purifying vWF which involves providing vWF-cp as a ligand, contacting a
XX solution comprising vWF with the polypeptide ligand under conditions
XX where vWF is bound to the ligand and recovering from the ligand purified
XX vWF. vWF-cp is useful for producing anti-vWF cp polypeptide antibodies
XX which involves immunising an animal with vWF-cp and isolating the anti-
XX vWF cp polypeptide antibodies from the animal. vWF-cp is useful for
XX producing a preparation of prophylaxis and therapy of thrombosis and
XX thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
XX Henoch-Schönlein purpura, preeclampsia, neonatal thrombocytopenia or
XX haemolytic-uraemic syndrome. vWF-cp can also be used for processing
XX plasmatic or recombinantly produced vWF. The invention is useful for
XX construction expression systems and generating transgenic animals which
XX express the polypeptide in vivo. The present sequence is human vWF-cp
XX fragment
XX
SQ Sequence 1353 AA;
Query Match 100.0%; Score 346; DB 5; Length 1353;
Best Local Similarity 100.0%; Pred. No. 4.2e-32;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 58
Db 1002 WHVGTWMECSVSCGDIQRRRDTCLGPOAQAPVPADFQCHLPKPKVTVRGWAGPCVQ 1059

RESULT 4
ID AAO16617 standard; protein; 1353 AA.
XX
XX
AC AAO16617;
```

```
XX 15-MAY-2003 (first entry)
DT Human von Willebrand factor (vWF)-cleaving enzyme-related protein #4.
DE
KW Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
KW von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
KW myocardial infarction; cerebral infarction; arteriosclerosis;
KW platelet thrombosis; stenosis.
XX Homo sapiens.
OS
XX WO200288366-A1.
PN
XX 07-NOV-2002.
PD
XX 25-APR-2002; 2002WO-JP004141.
PF
XX 25-APR-2001; 2001JP-00128342.
PR
XX 27-JUL-2001; 2001JP-00227510.
PR
XX 28-SEP-2001; 2001JP-00302977.
PR
XX 25-JAN-2002; 2002JP-00017596.
PR
XX (KAGA ) CHEMO-SERO-THERAPEUTIC RES INST.
PA
XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
XX WPI; 2003-120479/11.
PI
XX N-PSDB; ABT32583.
DR
XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
XX supplementary therapy for, thrombotic thrombocytopenic purpura, and for
XX developing drugs for e.g. myocardial infarction and cerebral infarction.
PT
XX Claim 4; Page 82-92; 144pp; Japanese.
PS
XX The invention comprises the amino acid and coding sequence of a von
XX Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
XX the invention are useful in the diagnosis and treatment of
XX thrombocytopenic purpura, and in developing drugs for myocardial
XX infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
XX and stenosis. The present amino acid sequence represents a human von
XX Willebrand factor (vWF)-cleaving enzyme-related protein
XX
XX Query Match 100.0%; Score 346; DB 6; Length 1353;
XX Best Local Similarity 100.0%; Pred. No. 4.2e-32;
XX Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WHVGTWMECSVSCGDGIQRRDTCLGPOAQPVPADFCQHLKPVTVRGCGWAGPCVQG 58
XX 1002 WHVGTWMECSVSCGDGIQRRDTCLGPOAQPVPADFCQHLKPVTVRGCGWAGPCVQG 1059
XX
XX RESULT 5
XX AAO16619
XX ID AAO16619 standard; protein; 1378 AA.
XX AC AAO16619;
XX
XX 15-MAY-2003 (first entry)
DT Human von Willebrand factor (vWF)-cleaving enzyme-related protein #6.
DE
XX Human; protease inhibitor; gene therapy; vWF-cleaving enzyme;
XX von Willebrand factor-cleaving enzyme; thrombocytopenic purpura;
XX myocardial infarction; cerebral infarction; arteriosclerosis;
XX platelet thrombosis; stenosis.
XX Homo sapiens.
OS
XX WO200288366-A1.
PN
```

```
XX 07-NOV-2002.
PD
XX 25-APR-2002; 2002WO-JP004141.
PF
XX 25-APR-2001; 2001JP-00128342.
PR
XX 27-JUL-2001; 2001JP-00227510.
PR
XX 28-SEP-2001; 2001JP-00302977.
PR
XX 25-JAN-2002; 2002JP-00017596.
PR
XX (KAGA ) CHEMO-SERO-THERAPEUTIC RES INST.
PA
XX Soejima K, Mimura N, Maeda H, Nozaki C, Hamamoto T, Nakagaki T;
XX WPI; 2003-120479/11.
PI
XX N-PSDB; ABT32585.
DR
XX von Willebrand factor-cleaving enzyme, applicable in diagnosis of, and
XX supplementary therapy for, thrombotic thrombocytopenic purpura, and for
XX developing drugs for e.g. myocardial infarction and cerebral infarction.
PT
XX Claim 4; Page 102-111; 144pp; Japanese.
PS
XX The invention comprises the amino acid and coding sequence of a von
XX Willebrand factor (vWF)-cleaving enzyme. The DNA and protein sequences of
XX the invention are useful in the diagnosis and treatment of
XX thrombocytopenic purpura, and in developing drugs for myocardial
XX infarction, cerebral infarction, arteriosclerosis, platelet thrombosis,
XX and stenosis. The present amino acid sequence represents a human von
XX Willebrand factor (vWF)-cleaving enzyme-related protein
XX
XX Query Match 100.0%; Score 346; DB 6; Length 1378;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-32;
XX Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 WHVGTWMECSVSCGDGIQRRDTCLGPOAQPVPADFCQHLKPVTVRGCGWAGPCVQG 58
XX 1027 WHVGTWMECSVSCGDGIQRRDTCLGPOAQPVPADFCQHLKPVTVRGCGWAGPCVQG 1084
XX
XX RESULT 6
XX ABB04153
XX ID ABB04153 standard; protein; 1416 AA.
XX AC ABB04153;
XX
XX 26-MAR-2002 (first entry)
DT Human ADAMTS-M polypeptide.
XX
XX Osteoarthritis; rheumatoid arthritis; inflammatory bowel disease;
XX Crohn's disease; asthma; Alzheimer's disease; organ transplant rejection;
XX cachexia; allergy; cancer; leukaemia; lymphoma; osteoporosis;
XX atherosclerosis; congestive heart failure; myocardial infarction; stroke;
XX neurodegenerative disease; autoimmune disorder; Huntington's;
XX Parkinson's; migraine; pain; depression; multiple sclerosis; burn;
XX infertility; diabetic shock; gene therapy; ADAMTS-M;
XX A Disintegrin And Metalloprotease; thrombospondin domain.
XX
XX Homo sapiens.
OS
XX
XX Location/Qualifiers
XX Key 1
XX Peptide /label= Prodomain
XX /note= "The mature form of the ADAMTS-M protein is
XX processed by furin cleavage of the prodomain"
XX 94..97
XX Cleavage-site /label= Furin_cleavage_site
XX 98..1416
XX Protein /label= Mature ADAMTS-M protein
XX /note= "The mature form of the ADAMTS-M protein is
XX processed by furin cleavage of the prodomain"
```

RESULT 7

CC The invention relates to an isolated or substantially pure Von Willebrand
 CC factor-cleaving protease (VWF-cp) polypeptide. VWF-cp is useful for
 CC purifying VWF which involves providing VWF-cp as a ligand, contacting a
 CC solution comprising VWF with the polypeptide ligand under conditions
 CC where VWF is bound to the ligand and recovering from the ligand purified
 CC VWF. VWF-cp is useful for producing anti-VWF cp polypeptide antibodies
 CC which involves immunising an animal with VWF-cp and isolating the anti-
 CC VWF cp polypeptide antibodies from the animal. VWF-cp is useful for
 CC producing a preparation of prophylaxis and therapy of thrombosis and
 CC thromboembolic disease such as thrombotic thrombocytopenic purpura (TTP),
 CC Henoch-Schönlein purpura, preeclampsia, neonatal thrombocytopenia or
 CC haemolytic-uraemic syndrome. VWF-cp can also be used for processing
 CC plasmatic or recombinantly produced VWF. The invention is useful for
 CC construction expression systems and generating transgenic animals which
 CC express the polypeptide in vivo. The present sequence is human VWF-cp
 CC protein. VWF-cp gene is located on chromosome 9
 XX
 XX
 XX Sequence 1427 AA;

Query Match 100.0%; Score 346; DB 5; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 4.5e-32;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WHVGTWMECSVSGDGIQRRDTCGLGPAQAPVPADFCQHLKPVTVRGCGWAGPCVGQ 58
 Db 1076 WHVGTWMECSVSGDGIQRRDTCGLGPAQAPVPADFCQHLKPVTVRGCGWAGPCVGQ 1133

RESULT 8
 ADD94038
 ID ADD94038 standard; protein; 1427 AA.
 XX
 XX ADD94038;
 XX
 XX 29-JAN-2004 (first entry)
 DT Human aggrecanase protein amino acid sequence.
 DE
 DE aggrecanase; aggrecan; articular cartilage; proteoglycan; proteolytic;
 KW cartilage degradation; osteoarthritis; inflammatory joint disease;
 KW antiarthritic; osteopathic; antiinflammatory;
 KW aggrecanase-associated disorder; osteoarthritis; inflammatory condition;
 KW human; enzyme.
 XX
 XX Homo sapiens.
 OS
 OS US2003105313-A1.
 PN
 XX 05-JUN-2003.
 PD
 XX 25-JAN-2002; 2002US-00057487.
 PF
 XX 16-OCT-2001; 2001US-00978979.
 PR
 XX (AMHP) AMERICAN HOME PROD CORP.
 PA
 XX Racie LA, Twine NC, Agostino MJ, Wolfman N, Morris EA;
 PI
 XX WFL; 2003-801251/75.
 DR N-PSDB; ADD94037.
 XX
 XX New isolated DNA molecule encoding an aggrecanase polypeptide for
 PT producing a purified human aggrecanase protein which can be used to
 PT develop inhibitors of aggrecanase.
 PT
 XX Claim 14; SEQ ID NO 8; 24pp; English.
 PS
 XX

CC This invention relates to a novel human aggrecanase protein and the DNA
 CC sequence which encodes it. Aggrecan is a major extracellular component of
 CC articular cartilage. It is a proteoglycan responsible for providing
 CC cartilage with its mechanical properties of compressibility and
 CC elasticity. A proteolytic activity (aggrecanase) is responsible for the
 CC cleavage of aggrecan thereby having a role in cartilage degradation
 CC

CC associated with osteoarthritis and inflammatory joint disease. Compounds
 CC which inhibit the activity of the protein of the invention may have
 CC antiarthritic, osteopathic or antiinflammatory activity. The invention
 CC may be used to produce a purified human aggrecanase protein. The protein
 CC (or fragment) may be used to develop inhibitors of aggrecanase, using
 CC three dimensional structural analysis or computer aided drug design. A
 CC peptide which binds to aggrecanase is used to inhibit the proteolytic
 CC degradation of aggrecan. The invention may be useful for the development
 CC of therapeutics for the treatment of aggrecanase-associated disorders,
 CC such as, osteoarthritis and other inflammatory conditions. The present
 CC sequence is that of the human aggrecanase protein (full length sequence)
 CC of the invention.
 XX
 XX Sequence 1427 AA;

Query Match 100.0%; Score 346; DB 7; Length 1427;
 Best Local Similarity 100.0%; Pred. No. 4.5e-32;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 WHVGTWMECSVSGDGIQRRDTCGLGPAQAPVPADFCQHLKPVTVRGCGWAGPCVGQ 58
 Db 1076 WHVGTWMECSVSGDGIQRRDTCGLGPAQAPVPADFCQHLKPVTVRGCGWAGPCVGQ 1133

RESULT 9
 ABB98125
 ID ABB98125 standard; protein; 1445 AA.
 XX
 XX ABB98125;
 XX
 XX 17-OCT-2002 (first entry)
 DT Human PMW Incyte ID 7473607CD1.
 DE
 DE
 XX
 XX Human; PMW; protein modification and maintenance molecule;
 KW anticonvulsant; neuroprotective; nootropic; cytostatic; antipsoriatic;
 KW antiasthmatic; dermatological; antidiabetic; antiparkinsonian;
 KW antianemic; antiinflammatory; antiulcer; antianginal; cardiac;
 KW hepatotropic; osteopathic; antiemetic; antipyretic; virucide;
 KW antibacterial; fungicide; gastrointestinal; antidiarrheic; laxative;
 KW haemostatic; anti-HIV; thrombolytic; anticoagulant; gynaecological;
 KW cardiovascular; antiarteriosclerotic; hypotensive; vasotropic;
 KW antitumour; antirheumatic; immunosuppressive; anti-allergic; antithyroid;
 KW nephrotropic; antitoxic; thyronimetic; antiarthritic; uropathic;
 KW ophthalmological; antiparasitic; tranquiliser; vulnerary; keratolytic;
 KW auditory; antiseborrheic; antidepressant; neuroleptic; antifertility;
 KW antihelmintic; protozoacide; Crohn's disease; hypertension; autoimmunity;
 KW inflammatory; anaemia; cell proliferative; developmental; epithelial;
 KW scabies; neurological; Alzheimer's disease; reproductive;
 KW ectopic pregnancy; gene therapy; vaccine; disorder;
 KW procollagen I N-proteinase.

OS Homo sapiens.
 XX
 XX WO200246383-A2.
 PN
 XX 13-JUN-2002.
 PD
 XX 05-DEC-2001; 2001WO-US046964.
 XX
 XX 08-DEC-2000; 2000US-0254399P.
 PR 21-DEC-2000; 2000US-0257803P.
 PR 05-JAN-2001; 2001US-0260110P.
 PR 19-JAN-2001; 2001US-0262851P.
 PR 25-JAN-2001; 2001US-0264623P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.

XX Yue H, Azimzai Y, Kallick DA, Baughn MR, Griffin JA, Swarnakar A;
 PI Lal PG, Walia NK, Hafalia AJA, Gandhi AR, Au-Young J, Elliott VS;
 PI Runkumar J, Thangavelu K, Lu Y, Warren BA, Lu DAM, Lee EA;
 PI Tribouley CM, Arvizu C, Delegeane AM, Yao MG, Khan FA;
 PI Sanjanwala MW;

XX WPI; 2002-519664/55.
DR N-PSDB; ABQ75946.
XX
PT New isolated Protein Modification and Maintenance polypeptides, useful
PT for diagnosis, and treatment of e.g. gastrointestinal disorders.
XX
PS Claim 1 (a); Page 154-157; 200pp; English.
XX
CC The invention relates to an isolated Protein Modification and Maintenance
CC (PMW) polypeptide. Polypeptides of the invention may be used in the
CC diagnosis, treatment and prevention of disorders associated with
CC decreased expression or activity of PMW. These include gastrointestinal
CC disorders (e.g. Crohn's disease), cardiovascular disorders (e.g.
CC hypertension), autoimmune/inflammatory disorders (e.g. anaemia), cell
CC proliferative disorders, developmental disorders, epithelial disorders
CC (e.g. scabies), neurological disorders (e.g. Alzheimer's disease)
CC reproductive disorders (e.g. ectopic pregnancy), and in gene therapy or a
CC vaccine for such diseases. They may also be used in the assessment of the
CC effects of exogenous compound on the expression of nucleic acid and amino
CC acid sequences of protein modification and maintenance molecules. The
CC current sequence represents a human PMW of the invention, which has been
CC found to have homology with B. taurus procollagen I N-proteinase
XX
XX Sequence 1445 AA;
SQ
Query Match 100.0%; Score 346; DB 5; Length 1445;
Best Local Similarity 100.0%; Pred. No. 4.6e-32;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WHVGTWMECSVSCDGIQRRDTCGLGPOAQPVPADFCQHLPKPVTVRGCGWAGPCVGQ 58
DB 1079 WHVGTWMECSVSCDGIQRRDTCGLGPOAQPVPADFCQHLPKPVTVRGCGWAGPCVGQ 1136
RESULT 10
ABB06011
ID ABB06011 standard; protein; 237 AA.
XX
AC ABB06011;
XX
DT 09-MAY-2002 (first entry)
XX
DE Human cancer suppressing protein PP1396 SEQ ID NO:41.
XX
KW Human; cancer suppression; cancer.
XX
OS Homo sapiens.
XX
PN CN133318-A.
XX
PD 19-SEP-2001.
XX
PF 14-MAR-2000; 2000CN-00111997.
XX
XX 14-MAR-2000; 2000CN-00111997.
PR
XX (SHAN-) SHANGHAI INST ONCOLOGY.
XX
PI Gu J, Yang S;
XX
XX WPI; 2002-042196/06.
DR N-PSDB; ABL39597.
XX
XX New human protein able to suppress growth of cancer cells and its
PT encoding polynucleotide.
PT
XX
PS Claim 1; Page 53 (Disclosure); 65pp; Chinese.
XX
CC The present invention describes human proteins with cancer suppressing
CC activity. Also described are the polynucleotides encoding the proteins
CC and a process for preparing the proteins by DNA recombination. The
CC proteins and polynucleotides can be used in the treatment of diseases

CC such as cancer. The present sequence represents a human cancer
CC suppressing protein from the present invention
XX
SQ Sequence 237 AA;
Query Match 35.4%; Score 122.5; DB 5; Length 237;
Best Local Similarity 41.0%; Pred. No. 2.3e-06;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;
QY 1 WHVGTWMECSVSCDGIQRRDTC---LGPQAQPVPADFCQHLPKPVTVRGCGWAGPCVG 57
DB 77 WYGTWMECSVSCDGIQRRDTCVSKLGTFFVTSPN-CSHLRPPALQ-----PCQG 130
QY 58 Q 58
DB 131 Q 131
RESULT 11
ADE09952
ID ADE09952 standard; protein; 358 AA.
XX
AC ADE09952;
XX
DT 29-JAN-2004 (first entry)
XX
DE Novel protein-related contig polypeptide sequence #540.
XX
KW novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; genetic disorder; contig.
XX
OS Unidentified.
XX
PN WC2003054152-A2.
XX
PD 03-JUL-2003.
XX
PF 10-DEC-2002; 2002WO-US039555.
XX
PR 10-DEC-2001; 2001US-0339739P.
PR 11-DEC-2001; 2001US-0339453P.
PR 14-MAR-2002; 2002US-0365091P.
PR 14-MAR-2002; 2002US-0365384P.
PR 12-APR-2002; 2002US-0372381P.
PR 12-APR-2002; 2002US-0372615P.
PR 22-APR-2002; 2002US-00128558.
PR 24-APR-2002; 2002US-0376045P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX
XX WPI; 2003-569235/53.
XX
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX
XX Disclosure; SEQ ID NO 3018; 1177pp; English.
XX
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence was used in the
CC exemplification of the invention.
XX
XX Sequence 358 AA;
SQ

```
Query Match      35.4%; Score 122.5; DB 7; Length 358;
Best Local Similarity 41.0%; Pred. No. 3.6e-06;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY      1 WHVGTWMECSVSCGDGIQRRTDC---LGPQAQAPVPADFQCHLPKPVTVRGWAGPCVG 57
Db      180 WITGPGECSECGSGTQRDIICVSKLTGTEFNVTSPSN-CSHLPRPDLQ-----PCQG 233
QY      58 Q 58
Db      234 Q 234

RESULT 12
AAE09698
ID AAE09698 standard; protein; 372 AA.
XX
AC AAE09698;
XX
DT 22-NOV-2001 (first entry)
XX
DE Human gene 9 encoding novel protein HUCMO06, SEQ ID NO:45.
XX
KW Human; cytostatic; gene therapy; inflammatory disorder; neural disorder;
KW Parkinson's disease; Alzheimer's disease; immune system disorder; AIDS;
KW autoimmune disease; rheumatoid arthritis; muscular disorder; ischaemia;
KW reproductive disorder; Crohn's disease; pulmonary disorder; cancer;
KW myocardial infarction; glomerulonephritis; nephrotic syndrome; tumour;
KW haematopoietic disorder; rhinitis; asthma; diabetes; atherosclerosis;
KW anti-sense therapy; endocrine disorder; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200155202-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001325.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198112P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0231415P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
```


PR	17-NOV-2000;	2000US-0243211P;
PR	17-NOV-2000;	2000US-0243211P;
PR	17-NOV-2000;	2000US-0243211P;
PR	17-NOV-2000;	2000US-0243213P;
PR	17-NOV-2000;	2000US-0243214P;
PR	17-NOV-2000;	2000US-0243215P;
PR	17-NOV-2000;	2000US-0243216P;
PR	17-NOV-2000;	2000US-0243217P;
PR	17-NOV-2000;	2000US-0243219P;
PR	17-NOV-2000;	2000US-0243244P;
PR	17-NOV-2000;	2000US-0243245P;
PR	17-NOV-2000;	2000US-0243264P;
PR	17-NOV-2000;	2000US-0243265P;
PR	17-NOV-2000;	2000US-0243297P;
PR	17-NOV-2000;	2000US-0243299P;
PR	17-NOV-2000;	2000US-0243300P;
PR	01-DEC-2000;	2000US-0250160P;
PR	01-DEC-2000;	2000US-0250391P;
PR	03-DEC-2000;	2000US-0251030P;
PR	03-DEC-2000;	2000US-0251988P;
PR	05-DEC-2000;	2000US-0256719P;
PR	05-DEC-2000;	2000US-0256719P;
PR	08-DEC-2000;	2000US-0251856P;
PR	08-DEC-2000;	2000US-0251868P;
PR	08-DEC-2000;	2000US-0251869P;
PR	08-DEC-2000;	2000US-0251989P;
PR	08-DEC-2000;	2000US-0251990P;
PR	11-DEC-2000;	2000US-0254097P;
PR	13-JAN-2001;	2001US-0259678P;
PR	17-JAN-2001;	2001US-0259630P;

Rosen CA, Ruben SM, Barash SC;

N-PSDB; ADC22050.

treating disorders

INVESTMENT

The invention relates to an isolated nucleic acid molecule (cDNA)

encoding a human secreted protein, representing one or more novel genes. Also included are recombinant vectors, host cells (expressing the protein), the secreted proteins (including their fragments, epitopes and homologues), an isolated antibody that binds specifically to the protein, diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or absence of a mutation in the nucleic acid and diagnosing a condition based on the presence or absence of the mutation), diagnosing a pathological condition or susceptibility to a pathological condition (comprising determining the presence or amount of expression of the protein in a biological sample and diagnosing a condition based on the presence or amount of expression of the protein), preventing, treating or ameliorating a medical condition of the protein), administering the nucleic acid or protein to a mammalian subject, by administering the nucleic acid or protein to the protein, the gene corresponding to identifying a binding partner to the protein, the gene corresponding to the cDNA sequence, and identifying an activity in a biological assay (comprising expressing the nucleic acid in a cell, isolating the supernatant, detecting an activity in a biological assay and identifying the protein in the supernatant having the activity). The nucleic acids and proteins display the following activities: Cytostatic, antibacterial, Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-

Query Match 35.4%: Score 122.5; DB 7; Length 372;

Best Local Similarity 41.0%; Pred. No. 3.7e-06;

Matches	25;	Conservative	7;	Mismatches	20
---------	-----	--------------	----	------------	----

1 WHYGTWMECSVSCGDIQRRDTC---LGPOAQPVPADFCQHLPKPVTVRGCGWAGPCVG 57

Db 212 WYTGPGECSECGSGTQRDRDIICVSKLGTGFNVTSPN-CSHLPRPALQ-----PCQG 265

QY 58 Q 58

Db 266 Q 266

RESULT 14

AAB50936

ID AAB50936 standard; protein; 491 AA.

AC AAB50936;

XX 20-MAR-2001 (first entry)

XX ADAM protein #2.

DE ADAM; disintegrin; metalloprotease; food additive; breast cancer;

KW ovarian.

XX Homo sapiens.

OS WO200073323-A2.

PN 07-DEC-2000.

PD 25-MAY-2000; 2000WO-US014308.

PF 27-MAY-1999; 99US-0136388P.

PR 09-JUL-1999; 99US-0142930P.

PR 28-JAN-2000; 2000US-0178717P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Wei P, Ni J, Hastings GA, Shi Y;

PI WPI; 2001-016507/02.

XX Seven nucleic acid molecules encoding ADAM polypeptides containing a

PT disintegrin and metalloprotease domain, useful in the prevention,

PT treatment and diagnosis of cancer, immune disorders, cardiovascular

PT disorders and neurological diseases.

XX Claim 11; Page 274-276; 287pp; English.

PS The present invention relates to seven members of the ADAMs (proteins

CC which contain A Disintegrin And Metalloprotease domain) protein family.

CC The ADAMs proteins and DNA may be used to treat disease, as a food

CC additive or preservative, for chromosome identification, as probes for

CC diagnosing a disorder related to the female reproductive system,

CC particularly breast and/or ovary cancer. They are also useful in the gene

CC therapy of breast and ovarian cancer

XX Sequence 491 AA;

SS Query Match 35.4%; Score 122.5; DB 4; Length 491;

CC Best Local Similarity 41.0%; Pred No 5e-06;

CC Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSGDGIQRREDTC---LGPQQAQVPVADFCQHLPRKPVTVRGWAGPCVG 57

Db 331 WYTGPGECSECGSGTQRDRDIICVSKLGTGFNVTSPN-CSHLPRPALQ-----PCQG 384

QY 58 Q 58

Db 385 Q 385

RESULT 15

ABU06026

ID ABU06026 standard; protein; 491 AA.

XX ABU06026;

AC

XX 02-MAY-2003 (first entry)

XX Human novel ADAM protein 2.

XX Human; enzyme; ADAM; a dystrophin and metalloprotease; cancer; stroke;

KW neural disease; cardiovascular disease; heart disease; atherosclerosis;

KW angina; reproductive disease; gastrointestinal disease; Crohn's disease;

KW endocrine disease; renal disease; central nervous system disease; AIDS;

KW Alzheimer's disease; Parkinson's disease; Huntington's disease; asthma;

KW schizophrenia; inflammatory disease; immunological disease; osteoporosis;

KW rheumatoid arthritis; inflammatory bowel disease; bone formation disease;

KW musculo-skeletal disease; angiogenic disease; blood related disorder;

XX respiratory neoplasia; thrombosis; gene therapy.

OS Homo sapiens.

XX US2002173640-A1.

XX 21-NOV-2002.

PD 19-APR-2002; 2002US-00125452.

PF 27-MAY-1999; 99US-0136388P.

PR 09-JUL-1999; 99US-0142930P.

PR 28-JAN-2000; 2000US-0178717P.

PR 25-MAY-2000; 2000WO-US014308.

PR 21-SEP-2000; 2000US-0234222P.

PR 16-NOV-2000; 2000US-00712907.

PR 19-SEP-2001; 2001US-00955504.

XX (HUMA-) HUMAN GENOME SCI INC.

PA Ruben SM, Ni J, Hastings GA, Shi Y, Wei P;

PI WPI; 2003-255274/25.

XX N-PSDB; ABX30527.

DR New isolated ADAM (a disintegrin and metalloprotease domain) polypeptide

XX PT and encoding polynucleotide useful for diagnosing and treating ADAM-

PT related pathologies e.g. cancer, blood disorders, inflammatory and immune

PT disorders and thrombosis.

XX Claim 11; Page 127-128; 147pp; English.

PS The invention relates to a new isolated polypeptide, comprising a

CC sequence at least 95% identical to any one of 7 fully defined sequences

CC given in the specification, or its fragment, epitope or variant. The

CC polypeptides are ADAM proteins (a dystrophin and metalloprotease). Also

CC included are the nucleic acids encoding the ADAM proteins, a vector

CC comprising the isolated nucleic acid, a host cell comprising the vector,

CC a recombinant host cell comprising the nucleic acid operably linked to a

CC heterologous regulating element which controls gene expression, an

CC isolated antibody that specifically binds to the ADAM protein.

CC identifying a binding partner to ADAM (comprising contacting ADAM with a

CC binding partner, and determining whether the binding partner affects an

CC activity of the polypeptide) and screening for molecules which modify

CC activities of ADAM (comprising contacting the polypeptide with a compound

CC suspected of having agonist or antagonist activity, and assaying for

CC activity of the polypeptide). The methods and compositions of the present

CC invention are useful for diagnosing, preventing, ameliorating and/or

CC treating diseases associated with the aberrant expression or activity of

CC the ADAM polypeptide, such as neural, cardiovascular (e.g. heart disease,

CC atherosclerosis, stroke, angina), reproductive, gastrointestinal (e.g.

CC Crohn's disease), endocrine, renal, central nervous system (e.g.

CC Alzheimer's, Parkinson's, Huntington's, schizophrenia), inflammatory,

CC immunological (e.g. AIDS, rheumatoid arthritis, asthma, inflammatory

CC bowel disease) bone formation, musculo-skeletal (e.g. osteoporosis),

CC angiogenic and blood related disorders, and respiratory neoplasias,

CC cancer and thrombosis. Many more diseases and disorders are listed in the

CC specification. They can also be used in chromosome identification,

CC differential identification of tissue and cell types, screening assays

CC and molecular weight markers. The present sequence represents an ADAM

CC protein
XX
SQ Sequence 491 AA;
Query Match 35.4%; Score 122.5; DB 6; Length 491;
Best Local Similarity 41.0%; Pred. No. 5e-06;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;
QY 1 WHVGTWMECSVSCGDGIQRRTDC---LGPQAQAPVPADFCQHLPKPVTYRGWAGPCVG 57
Db 331 WYTGPMGECSSGSGTQRDIIICVSKLGTFFNVTSPSN-CSHLPRPPALQ-----PCQG 384
QY 58 Q 58
Db 385 Q 385

Search completed: March 13, 2004, 07:39:07
Job time : 9.48854 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:43:25 ; Search time 4.59167 Seconds
(without alignments)
2667.199 Million cell updates/sec

Title: US-09-836-712-2_COPY_1099_1156
Perfect score: 346
Sequence: 1 WHVGTWMECSVSCGDIQRR.....QHLPKPVTVRGCGWAGPCVCGQ 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues
Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	346	100.0	1416	9	US-09-836-712-2
2	346	100.0	1427	14	US-10-222-334-2
3	346	100.0	1427	14	US-10-057-487-8
4	122.5	35.4	372	9	US-09-764-903-45
5	122.5	35.4	491	9	US-09-955-504-10
6	122.5	35.4	491	13	US-10-135-470-10
7	122.5	35.4	491	13	US-10-135-452-10
8	122.5	35.4	645	5	US-09-764-853-762
9	121.5	35.1	56	14	US-10-087-887-89
10	121.5	35.1	283	14	US-10-087-887-16
11	121.5	35.1	367	9	US-09-764-903-45
12	121.5	35.1	369	9	US-09-764-903-43
13	121.5	35.1	1081	15	US-10-120-801-46
14	121.5	35.1	1123	15	US-10-120-801-45
15	121.5	35.1	1162	15	US-10-161-493-124

16	121.5	35.1	1221	14	US-10-240-545A-2
17	121.5	35.1	1745	9	US-09-800-729-89
18	121.5	35.1	1745	11	US-09-833-245-2222
19	121.5	35.1	1762	13	US-10-044-807-2
20	121.5	35.1	1766	9	US-09-764-853-554
21	119.5	34.5	1629	9	US-09-972-467-2
22	119.5	34.5	1882	9	US-09-918-171A-13
23	119.5	34.5	1907	9	US-09-938-330-25
24	119.5	34.5	1916	15	US-10-274-639-10
25	118	34.1	1686	15	US-10-386-414-2
26	118	34.1	1690	9	US-09-788-043C-5
27	116.5	33.7	755	15	US-10-093-463-30
28	116.5	33.7	950	9	US-09-965-631-4
29	116	33.5	1280	14	US-10-087-887-86
30	115.5	33.4	481	13	US-10-105-929-8
31	115.5	33.4	481	14	US-10-365-227-8
32	115	33.2	1023	10	US-09-789-390-25
33	115	33.2	1023	10	US-09-789-390-27
34	115	33.2	1023	10	US-09-789-390-28
35	115	33.2	1023	10	US-09-789-390-59
36	115	33.2	1023	10	US-09-789-390-61
37	115	33.2	1281	10	US-09-789-390-2
38	115	33.2	1281	10	US-09-789-390-5
39	115	33.2	1617	9	US-09-784-358-16
40	115	33.2	1691	9	US-09-784-358-2
41	115	33.2	1691	10	US-09-789-390-4
42	114	32.9	771	9	US-09-784-358-14
43	114	32.9	845	9	US-09-784-358-12
44	114	32.9	1311	14	US-10-103-377C-6
45	113.5	32.8	874	10	US-09-796-753-20

ALIGNMENTS

RESULT 1

US-09-836-712-2
; Sequence 2, Application US/09836712
; Patent No. US20010049106A1
; GENERAL INFORMATION:
; APPLICANT: PFIZER INC.
; TITLE OF INVENTION: ADAMTS POLYPEPTIDES, NUCLEIC ACIDS ENCODING THEM, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: PCT0851A
; CURRENT APPLICATION NUMBER: US/09/836,712
; CURRENT FILING DATE: 2001-04-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1416
; TYPE: PRT
; ORGANISM: Human
US-09-836-712-2

Query Match 100.0%; Score 346; DB 9; Length 1416;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCOHLPKPVTVRGCGWAGPCVCGQ 58
Db	1099	WHVGTWMECSVSCGDIQRRDTCIGPQAQAPVPADFCOHLPKPVTVRGCGWAGPCVCGQ 1156

RESULT 2

US-10-222-334-2
; Sequence 2, Application US/10222334
; Publication No. US20030073116A1
; GENERAL INFORMATION:
; APPLICANT: Ginsburg, David
; APPLICANT: Levy, Gallia
; APPLICANT: Tsai, Han-Mou
; TITLE OF INVENTION: ADAMTS13 Genes and Proteins and Variants, and Uses Thereof
; FILE REFERENCE: UN-07288

```
; CURRENT APPLICATION NUMBER: US/10/222,334
; CURRENT FILING DATE: 2002-08-16
; PRIOR APPLICATION NUMBER: 60/312,934
; PRIOR FILING DATE: 2001-08-16
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-222-334-2

Query Match      100.0%; Score 346; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 58
   |||||
Db 1076 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 1133
   |||||

RESULT 3
US-10-057-487-8
; Sequence 8, Application US/10057487
; Publication No. US20030105313A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; TITLE OF INVENTION: Aggrecanase Molecules
; FILE REFERENCE: 08702.0073
; CURRENT APPLICATION NUMBER: US/10/057,487
; CURRENT FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 60/241,469
; PRIOR FILING DATE: 2000-10-18
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 1427
; TYPE: PRT
; ORGANISM: homo sapiens
US-10-057-487-8

Query Match      100.0%; Score 346; DB 14; Length 1427;
Best Local Similarity 100.0%; Pred. No. 2.5e-30;
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 58
   |||||
Db 1076 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 1133
   |||||

RESULT 4
US-09-764-903-45
; Sequence 45, Application US/09764903
; Patent No. US20020090674A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT228
; CURRENT APPLICATION NUMBER: US/09/764,903
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-903-45

Query Match      35.4%; Score 122.5; DB 9; Length 372;
Best Local Similarity 41.0%; Pred. No. 8.4e-06;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 57
   |||||
Db 331 WYTGPWGECSSECGSGTQRDIICVSKLGTFFNVTSPSN-CSHLPRPPALQ-----PCQG 384
   |||||

RESULT 5
US-09-955-504-10
; Sequence 10, Application US/09955504
; Publication No. US20020182702A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT006P2
; CURRENT APPLICATION NUMBER: US/09/955,504
; CURRENT FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/234,222
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 09/712,907
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: PCT/US00/14308
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/178,717
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/142,930
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/136,388
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-955-504-10

Query Match      35.4%; Score 122.5; DB 9; Length 491;
Best Local Similarity 41.0%; Pred. No. 1.1e-05;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCGDIQRRDTC LGPQAQAPVPADFCQHLKPVTVRGCGWAGPCVG 57
   |||||
Db 331 WYTGPWGECSSECGSGTQRDIICVSKLGTFFNVTSPSN-CSHLPRPPALQ-----PCQG 384
   |||||

RESULT 6
US-10-125-470-10
; Sequence 10, Application US/10125470
; Publication No. US20020165377A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PT006P1
; CURRENT APPLICATION NUMBER: US/10/125,470
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US/09/712,907A
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/136,388
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/142,930
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/178,717
; PRIOR FILING DATE: 2000-01-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
```

```
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-470-10

Query Match      35.4%; Score 122.5; DB 13; Length 491;
Best Local Similarity 41.0%; Pred. No. 1.1e-05;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCDGIQRDRDTC---LGPOAQAPVPADFCOHLPKPVTVRGCGWAGPCVG 57
Db 331 WYTGPGWCESECGSGTQRDDIICVSKLGTFTNVTSPN-CSHLPRPALQ-----PCQG 384

QY 58 Q 58
Db 385 Q 385

RESULT 7
US-10-125-452-10
; Sequence 10, Application US/10125452
; Publication No. US20020173640A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: ADAM Polynucleotides, Polypeptides, and Antibodies
; FILE REFERENCE: PTO06P2
; CURRENT APPLICATION NUMBER: US/10/125.452
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/955,504
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 09/712,907
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: PCT/US00/14308
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/178,717
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/142,930
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 60/136,388
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 491
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-452-10

Query Match      35.4%; Score 122.5; DB 13; Length 491;
Best Local Similarity 41.0%; Pred. No. 1.1e-05;
Matches 25; Conservative 7; Mismatches 20; Indels 9; Gaps 3;

QY 1 WHVGTWMECSVSCDGIQRDRDTC---LGPOAQAPVPADFCOHLPKPVTVRGCGWAGPCVG 57
Db 331 WYTGPGWCESECGSGTQRDDIICVSKLGTFTNVTSPN-CSHLPRPALQ-----PCQG 384

QY 58 Q 58
Db 385 Q 385

RESULT 8
US-09-764-853-762
; Sequence 762, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PIZ06
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 762
; LENGTH: 645
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-853-762

Query Match      35.4%; Score 122.5; DB 9; Length 645;
Best Local Similarity 38.2%; Pred. No. 1.4e-05;
Matches 29; Conservative 6; Mismatches 22; Indels 19; Gaps 4;

QY 1 WHVGTWMECSVSCDGIQRDRDTCGLGPQA---QAPVPADFCOHLPK-PVTVRGC-----50
Db 432 WMVTSWACTRSCGGVQTRRVTCQKLKASGISFVSNMDMCTQVAKRPVDTQACNQQLCV 491

QY 51 -WA-----GPCVG 57
Db 492 EWAFSSWGQCNGPCVG 507

RESULT 9
US-10-087-887-89
; Sequence 89, Application US/10087887
; Publication No. US20030198957A1
; GENERAL INFORMATION:
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Conley, Pamela B.
; APPLICANT: Yang, Ruey-Bing
; APPLICANT: Hart, Matthew
; APPLICANT: Tomlinson, James E.
; APPLICANT: Topper, James N.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Leach, Martin D.
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Komuves, Laszlo
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-285
; CURRENT APPLICATION NUMBER: US/10/087,887
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/273,049
; PRIOR FILING DATE: 2001-03-02
; PRIOR APPLICATION NUMBER: 60/279,893
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/277,791
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/281,248
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/282,864
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,537
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/282,867
; PRIOR FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 89
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-887-89

Query Match      35.1%; Score 121.5; DB 14; Length 56;
Best Local Similarity 38.2%; Pred. No. 1.7e-06;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;

QY 1 WHVGTWMECSVSCDGIQRDRDTCGLGPQAQAPVPADFCOHLPKPVTVRGCGWAGPC 55
Db 3 WSLGFWSCSKTCGRGVKRELLCKGSAET-LPESQCTSLRPELQEGCVLGR 56

RESULT 10
US-10-087-887-16
```

```

, Sequence 16, Application US/10087887
, Publication No. US20030198957A1
, GENERAL INFORMATION:
, APPLICANT: Kekuda, Ramesh
, APPLICANT: Conley, Pamela B.
, APPLICANT: Yang, Ruey-Bing
, APPLICANT: Hart, Matthew
, APPLICANT: Tomlinson, James E.
, APPLICANT: Topper, James N.
, APPLICANT: Shminks, Richard A.
, APPLICANT: Leach, Martin D.
, APPLICANT: Zehusen, Bryan D.
, APPLICANT: Komuves, Laszlo
, APPLICANT: Padigattu, Muralidhara
, TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS
, FILE REFERENCE: 21402-285
, CURRENT APPLICATION NUMBER: US/10/0-03-01
, CURRENT FILING DATE: 2002-03-01
, PRIOR APPLICATION NUMBER: 60/273,041
, PRIOR FILING DATE: 2001-03-02
, PRIOR APPLICATION NUMBER: 60/279,881
, PRIOR FILING DATE: 2001-03-29
, PRIOR APPLICATION NUMBER: 60/277,799
, PRIOR FILING DATE: 2001-03-21
, PRIOR APPLICATION NUMBER: 60/281,241
, PRIOR FILING DATE: 2001-04-03
, PRIOR APPLICATION NUMBER: 60/282,861
, PRIOR FILING DATE: 2001-04-10
, PRIOR APPLICATION NUMBER: 60/282,531
, PRIOR FILING DATE: 2001-04-09
, PRIOR APPLICATION NUMBER: 60/282,861
, PRIOR FILING DATE: 2001-04-10
, NUMBER OF SEQ ID NOS: 104
, SOFTWARE: CuraseqList version 0.1
, SEQ ID NO 16
, LENGTH: 283
, TYPE: PRT
, ORGANISM: Homo sapiens
, US-10-087-887-16

```

```

Query Match      35.1%; Score 121.5; DB 14; Length 283;
Best Local Similarity 38.2%; Pred. No. 8.4e-06;
Matches 24; Conservative 9; Mismatches 24; Indels 1; Gaps 1;

QY      1  WHVCTWECVSVCSDGIQRRDTCLGQAOAPVPADFCQHLPKPVTYRGCGWAGPC 55
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB      57  WSLGFWSCSKTCGRGVRKRELLCKGSAAT-LPESQCTSLRPELOEGCVLGRG 110

RESULT 11
US-09-764-903-46
; Sequence 46, Application US/09764903
; Patent No. US20020090674A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ28
; CURRENT APPLICATION NUMBER: US/09/764.903
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 46
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (168)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (226)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

```

US-09-764-903-46

```
Query Match          35.1%; Score 121.5; DB 9; Length 367;
Best Local Similarity 38.2%; Pred.No. 1.1e-05;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;
```

RESULT 12

```

US-09-764-903-43
; Sequence 43, Application US/09764903
; Patent No. US20020090674A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PRZ28
; CURRENT APPLICATION NUMBER: US/09/764,903
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-903-43

```

US-09-764-903-43

Query Match 35.1%; Score 121.5; DB 9; Length 369;
Best Local Similarity 38.2%; Ref. No. 1.a-05;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;
QY 1 WHVGTWMECSVCGDGIQRRRTCLGFAQAQAPVADFQHLPKPVTYVRCWAGPC 55
p0 143 WSLGPGWSCKTGKGRVYKRELLCKSAAT-LPESOCTSLPRPELOECVLGRC 196

RESULT 13

RESUB: 13
US-10-120-801-46
; Sequence 46, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malvankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284

; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 1081
; TYPE: PRT
; ORGANISM: human
US-10-120-801-46

Query Match 35.1%; Score 121.5; DB 15; Length 1081;
Best Local Similarity 38.2%; Pred. No. 3.1e-05;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;

Qy 1 WHVGTWMECSVSCGDIQRRDTCILGPQAQAPVPADFQCHLPKPVTVRGWAGPC 55
Db 994 WSLGFWSCSKTCGRVRRKLLCKGSAABT-LPESQCTSLRPPELQEGCVLGR 1047

RESULT 14
US-10-120-801-45
; Sequence 45, Application US/10120801
; Publication No. US20030203843A1
; GENERAL INFORMATION:
; APPLICANT: Pena, Carol
; APPLICANT: Guo, Xiaojia
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Mehraban, Fuad
; APPLICANT: Topper, James N.
; APPLICANT: Malyankar, Uriel
; APPLICANT: Wasserman, Scott
; APPLICANT: Edinger, Shlomit
; APPLICANT: Smithson, Glennnda
; APPLICANT: Gunther, Erik
; APPLICANT: Komuves, Laszlo

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-340
; CURRENT APPLICATION NUMBER: US/10/120,801
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 60/285748
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 60/286068
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/286292
; PRIOR FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: 60/288334
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/291241
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 60/322284
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/285609
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 1123
; TYPE: PRT
; ORGANISM: human
US-10-120-801-45

Query Match 35.1%; Score 121.5; DB 15; Length 1123;
Best Local Similarity 38.2%; Pred. No. 3.2e-05;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;

Qy 1 WHVGTWMECSVSCGDIQRRDTCILGPQAQAPVPADFQCHLPKPVTVRGWAGPC 55
Db 861 WSLGFWSCSKTCGRVRRKLLCKGSAABT-LPESQCTSLRPPELQEGCVLGR 914

RESULT 15
US-10-161-493-124
; Sequence 124, Application US/10161493
; Publication No. US20040018555A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Li, Li
; APPLICANT: Zhong, Mei
; APPLICANT: Casman, Stacie J
; APPLICANT: Gerlach, Valerie
; APPLICANT: Shimkets, Richard A
; APPLICANT: Gorman, Linda
; APPLICANT: Pena, Carol BA
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Leite, Mario W
; APPLICANT: Rastelli, Luca
; APPLICANT: MacDougall, John R
; APPLICANT: Taupier Jr., Raymond J
; APPLICANT: Guo, Xiaojia Sasha
; APPLICANT: Miller, Charles E
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Hjal, Tord
; APPLICANT: Voss, Edward Z
; APPLICANT: Boldog, Ferenc L
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Ji, Weizhen
; APPLICANT: Smithson, Glennnda
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Millet, Isabelle
; APPLICANT: Ellerman, Karen
; TITLE OF INVENTION: No. US20040018555A1el Antibodies that Bind to Antigenic Polypep
; TITLE OF INVENTION: Acids Encoding the Antigens, and Methods of Use
; FILE REFERENCE: 21402-377A
; CURRENT APPLICATION NUMBER: US/10/161,493
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: 60/295,607
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: 60/337,524
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/296,404
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,418
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: 60/296,575
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: 60/359,151
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/297,414
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: 60/297,573
; PRIOR FILING DATE: 2001-06-12
; PRIOR APPLICATION NUMBER: 60/341,143
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: 60/297,567
; PRIOR FILING DATE: 2001-06-12
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 299
; SEQ ID NO 124
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-161-493-124

Query Match 35.1%; Score 121.5; DB 15; Length 1162;
Best Local Similarity 38.2%; Pred. No. 3.3e-05;
Matches 21; Conservative 9; Mismatches 24; Indels 1; Gaps 1;
Qy 1 WHVGTWMECSVSCGDIQRRDTCILGPQAQAPVPADFQCHLPKPVTVRGWAGPC 55

Db 936 WSLGPNQCSKTCGRGVKRELLCKGSAAT-LPESQCTSLPRPELQEGCVLGR 989

Search completed: March 13, 2004, 08:07:06
Job time : 5.59167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 13, 2004, 07:34:04 ; Search time 2.44688 Seconds
(without alignments)
1223.727 Million cell updates/sec

Title: US-09-836-712-2_COPY_1099_1156
Perfect score: 346
Sequence: 1 WHVGTWMECSVSCGDIQR...QHLPKPVTVRGWAGPCVQG 58

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	121.5	35.1	1745	US-09-800-729-89	Sequence 89, Appl
2	119.5	34.5	1882	US-09-369-364A-13	Sequence 13, Appl
3	115.5	33.4	481	US-09-130-491-8	Sequence 8, Appl
4	112	32.4	2165	US-09-800-729-155	Sequence 155, App
5	107.5	31.1	53	US-09-800-729-165	Sequence 165, App
6	106.5	30.8	1211	US-09-491-522-5	Sequence 5, Appl
7	103	29.8	59	US-09-800-729-171	Sequence 171, App
8	101.5	29.3	56	US-09-800-729-173	Sequence 173, App
9	101.5	29.3	1059	US-09-800-729-217	Sequence 217, App
10	100.5	29.0	905	US-09-369-364A-9	Sequence 9, Appl
11	95.5	27.6	58	US-09-800-729-169	Sequence 169, App
12	95.5	27.6	514	US-09-800-729-124	Sequence 124, App
13	95	27.5	60	US-09-800-729-167	Sequence 167, App
14	93.5	27.0	1224	US-09-930-872-4	Sequence 4, Appl
15	93	26.9	788	US-09-18-914-4	Sequence 4, Appl
16	92.5	26.7	551	US-09-130-491-16	Sequence 16, Appl
17	92.5	26.7	727	US-09-445-023A-12	Sequence 12, Appl
18	91.5	26.4	727	US-09-445-023A-1	Sequence 1, Appl
19	91.5	26.4	949	US-09-568-559-2	Sequence 2, Appl
20	91.5	26.4	967	US-09-130-491-2	Sequence 2, Appl
21	89.5	25.9	1205	US-09-491-522-11	Sequence 11, Appl
22	87	25.1	58	US-09-800-729-168	Sequence 168, App
23	85.5	24.7	61	US-09-800-729-166	Sequence 166, App
24	85	24.6	58	US-09-800-729-172	Sequence 172, App
25	85	24.6	525	US-09-369-364A-21	Sequence 21, Appl
26	84.5	24.4	1081	US-09-369-364A-17	Sequence 17, Appl
27	84	24.3	59	US-09-800-729-163	Sequence 163, App

ALIGNMENTS

RESULT 1
US-09-800-729-89
; Sequence 89, Application US/09800729
; Patent No. 6505592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044PI
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 89
; LENGTH: 1745
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-89

Query Match	35.1%	Score	121.5	DB	4	Length	1745
Best Local Similarity	36.8%	Pred. No.	6.6e-06				
Matches	28	Conservative	7	Mismatches	22	Indels	19
Gaps	4						
QY	1	WHVGTWMECSVSCGDIQRRTDCLGPQA--QAPVPADFQHLPK-PVTVRG	-----	50			
Db	1532	WMYTSKSACTRSCGGVQTRRTVCQKLKASGISPVSNMCTQVAKRPVDTQACNQQLCV	1591				
QY	51	-WA-----GPCVG	57				
Db	1592	EWAFSSWGQCGPCIG	1607				
RESULT 2							
US-09-369-364A-13							
; Sequence 13, Application US/09369364A							
; Patent No. 6391610							
; GENERAL INFORMATION:							
; APPLICANT: Apte, Suneel							
; APPLICANT: Hurskainen, Tiina L.							
; APPLICANT: Hirohata, Satoshi							
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases							
; FILE REFERENCE: 26473/4007/10-30-00							
; CURRENT APPLICATION NUMBER: US/09/369,364A							
; CURRENT FILING DATE: 1999-08-06							
; NUMBER OF SEQ ID NOS: 31							
; SOFTWARE: PatentIn Ver. 2.1							
; SEQ ID NO 13							

```

; LENGTH: 1882
; TYPE: PRT
; ORGANISM: Homo sapiens ADAMTS-9
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (468)
; OTHER INFORMATION: Xaa = C
; NAME/KEY: MOD_RES
; LOCATION: (521)
; OTHER INFORMATION: Xaa = Y
;
US-09-369-364A-13

Query Match          34.5%; Score 119.5; DB 4; Length 1882;
Best Local Similarity 40.0%; Pred. No. 1.2e-05;
Matches 22; Conservative 8; Mismatches 24; Indels 1; Gaps 1;

QY 1 WHVGTWMECSVSCGDIQRRDTCGLGPOAQAPVPADFCQHLKPVTVRGCGWAGPC 55
Db 1392 WSTGFWSSCSVSCGRGHQNRVYCWAKDG-SHLESYCKHLAKPHGRKCRGRC 1445

RESULT 3
US-09-130-491-8
; Sequence 8, Application US/09130491
; Patent No. 6416974
; GENERAL INFORMATION:
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: TANGO-71, TANGO-73, TANGO-74, TANGO-76, AND TANGO-83
; FILE REFERENCE: 09404/041001
; CURRENT APPLICATION NUMBER: US/09/130,491
; CURRENT FILING DATE: 1998-08-07
; EARLIER APPLICATION NUMBER: US 60/058,108
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: US 60/054,961
; EARLIER FILING DATE: 1997-08-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 481
; TYPE: PRT
; ORGANISM: Rattus rattus
;
US-09-130-491-8

Query Match          33.4%; Score 115.5; DB 4; Length 481;
Best Local Similarity 39.3%; Pred. No. 8.4e-06;
Matches 22; Conservative 6; Mismatches 25; Indels 3; Gaps 1;

QY 1 WHVGTWMECSVSCGDIQRRDTCGLGPOAQAPVPADFCQHLKPVTVRGCGWAGPCV 56
Db 428 WVLGDWSECPSTCGGGWQRRTVECRDPGQA---SDTCDEALKPEDAKPCGSPCL 480

RESULT 4
US-09-800-729-155
; Sequence 155, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 155
; LENGTH: 2165
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-800-729-155

Query Match          32.4%; Score 112; DB 4; Length 2165;
Best Local Similarity 38.2%; Pred. No. 0.00011;
Matches 21; Conservative 6; Mismatches 26; Indels 2; Gaps 1;

QY 1 WHVGTWMECSVSCGDIQRRDTCGLGPOAQAPVPADFCQHLKPVTVRGCGWAGPC 55
Db 1371 WATGPMTCATCGNGTQRLLKC--RDHVRDLPDEYCNHLDKEVSTNCRLRDC 1423

RESULT 5
US-09-800-729-165
; Sequence 165, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 165
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-800-729-165

Query Match          31.1%; Score 107.5; DB 4; Length 53;
Best Local Similarity 42.6%; Pred. No. 7e-06;
Matches 23; Conservative 7; Mismatches 19; Indels 5; Gaps 2;

QY 6 WMCSVSCGDIQRRDTC---LGPOAQAPVPADFCQHLKPVTVRGCGWAGPC 55
Db 1 WSACTVTCGVGTQVIVRCQVLLSFSQSVADLPIDCECG-PKPASQACVAGPC 53

RESULT 6
US-09-491-522-5
; Sequence 5, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Collige, Alain
; APPLICANT: Lepiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE
; TITLE OF INVENTION: AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/491,522
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/886,333
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;
```

```

; NAME: Abrams, Samuel B
; REGISTRATION NUMBER: 30,605
; REFERENCE/DOCKET NUMBER: 8389-0060-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1211 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-491-522-5

Query Match      30.8%; Score 106.5; DB 4; Length 1211;
Best Local Similarity 40.0%; Pred. No. 0.00026;
Matches 22; Conservative 4; Mismatches 24; Indels 5; Gaps 2;

QY 1 WHVGTWMECSVSCGDGIQRRRDTCLGPOAQAPVPADFQCHLPKPVTVRGCGWAGPC 55
Db 979 WRAGPWSQCSVTGNGTQTRPVLC---RTADDSFGICQE-ERPETARTCLRGPC 1028

RESULT 7
US-09-800-729-171
; Sequence 171, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 171
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-171

Query Match      29.8%; Score 103; DB 4; Length 59;
Best Local Similarity 40.7%; Pred. No. 2.7e-05;
Matches 22; Conservative 6; Mismatches 22; Indels 4; Gaps 2;

QY 1 WHVGTWMECSVSCGDGIQRRRDTCLGPOA---QAPVPADFQCHLPK-PVTVRGCG 50
Db 1 WMTVSWSACTRS CGGVQTRRVTCQKLKASGISTPVSNDMCTQVAKRPVDTQAC 54

RESULT 8
US-09-800-729-173
; Sequence 173, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 173
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-173

Query Match      29.3%; Score 101.5; DB 4; Length 56;
Best Local Similarity 33.9%; Pred. No. 3.8e-05;
Matches 19; Conservative 7; Mismatches 29; Indels 1; Gaps 1;

QY 1 WHVGTWMECSVSCGD-GIQRRTDTCGLGPOAQAPVPADFQCHLPKPVTVRGCGWAGPC 55
Db 1 WRVSLWTLCTATCGNIGFQSRREVCVARTNKAVPEHLCSWGPRPANWQRCNITPC 56

RESULT 9
US-09-800-729-217
; Sequence 217, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 217
; LENGTH: 1059
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-800-729-217

Query Match      29.3%; Score 101.5; DB 4; Length 1059;
Best Local Similarity 33.9%; Pred. No. 0.00087;
Matches 19; Conservative 7; Mismatches 21; Indels 9; Gaps 2;

QY 1 WHVGTWMECSVSCGDGIQRRRDTCLGPOAQAPVPADFQCHLPKPVTVRGCGWAG-PC 55
Db 296 WEAGKWSKCTASCQGVRRRHVACVG-----GSDCDEGGRPRQRTTCYAGIPC 343

RESULT 10
US-09-369-364A-9
; Sequence 9, Application US/09369364A
; Patent No. 6391610
; GENERAL INFORMATION:
; APPLICANT: Apte, Suneel
; APPLICANT: Hurskainen, Tiina L.
; APPLICANT: Hirohata, Satoshi
; TITLE OF INVENTION: Nucleic Acids Encoding Zinc Metalloproteases
; FILE REFERENCE: 26473/4007/10-30-00
; CURRENT APPLICATION NUMBER: US/09/369,364A
; CURRENT FILING DATE: 1999-08-06
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 905
; TYPE: PRT
; ORGANISM: Mus musculus ADAMTS-8
; US-09-369-364A-9

Query Match      29.0%; Score 100.5; DB 4; Length 905;
Best Local Similarity 36.4%; Pred. No. 0.00097;
Matches 20; Conservative 5; Mismatches 27; Indels 3; Gaps 1;

QY 1 WHVGTWMECSVSCGDGIQRRRDTCLGPOAQAPVPADFQCHLPKPVTVRGCGWAGPC 55
Db 852 WVLGDWSECPSTCRGSWQRRVTECRDPSCQA---SDTCDEALKPEDAKPCGSQPC 903
```

```
RESULT 11
US-09-800-729-169
; Sequence 169, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 169
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-169

Query Match      27.6%; Score 95.5; DB 4; Length 58;
Best Local Similarity 34.5%; Pred. No. 0.0002;
Matches 20; Conservative 5; Mismatches 30; Indels 3; Gaps 1;

Qy 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
Db 1 WLLSDWTECSVSCGEGTQTRSAICRKLTKLSTVNSTLCPLPSSSIRPCMCLATC 58

RESULT 12
US-09-800-729-124
; Sequence 124, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 124
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-124

Query Match      27.6%; Score 95.5; DB 4; Length 514;
Best Local Similarity 34.5%; Pred. No. 0.0021;
Matches 20; Conservative 5; Mismatches 30; Indels 3; Gaps 1;

Qy 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQAPVPADFCQHLKPVTVRGCGWAGPC 55
Db 83 WLLSDWTECSVSCGEGTQTRSAICRKLTKLSTVNSTLCPLPSSSIRPCMCLATC 140

RESULT 13
US-09-800-729-167
; Sequence 167, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: P2044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
```

```
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 167
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-800-729-167

Query Match      27.5%; Score 95; DB 4; Length 60;
Best Local Similarity 30.9%; Pred. No. 0.00024;
Matches 21; Conservative 5; Mismatches 10; Indels 32; Gaps 3;

Qy 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQAPVPADFCQHL-----P 42
Db 1 WEIGKSPCSLTCGVLQ-TRDV-----FCSHLLSREMNETVILADELCRQP 46

Qy 43 KPVTVRGC 50
Db 47 KPSTVOAC 54

RESULT 14
US-09-930-872-4
; Sequence 4, Application US/09930872
; Patent No. 6448388
; GENERAL INFORMATION:
; APPLICANT: Fridde, Carl Johan
; APPLICANT: Hilbun, Erin
; TITLE OF INVENTION: No. 648388e1 Human Proteases and Polynucleotides Encoding the:
; FILE REFERENCE: LEX-0219-USA
; CURRENT APPLICATION NUMBER: US/09/930,872
; CURRENT FILING DATE: 2001-08-14
; PRIOR APPLICATION NUMBER: US 60/225,852
; PRIOR FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1224
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-930-872-4

Query Match      27.0%; Score 93.5; DB 4; Length 1224;
Best Local Similarity 36.2%; Pred. No. 0.0089;
Matches 21; Conservative 5; Mismatches 29; Indels 3; Gaps 2;

Qy 1 WHVGTWMECSVSCGDIQRRDTC---LGPQAQAP-VPADFCQHLKPVTVRGCGWAGPC 55
Db 990 WSAGFMAECSTCGKMGKRAVACKSTNPSAQILLPAVCTSEPKPRVHEACLQRC 1047

RESULT 15
US-08-918-914-4
; Sequence 4, Application US/08918914
; Patent No. 5876963
; GENERAL INFORMATION:
; APPLICANT: Mitchell, Peter
; APPLICANT: Hutchinson, Nancy
; APPLICANT: Lawton, Michael
; APPLICANT: Magma, Holly
; APPLICANT: Vocum, Sue
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
```

```
;
;
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/918,914
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 788 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1070094
;
US-08-918-914-4

Query Match 26.9%; Score 93; DB 2; Length 788;
Best Local Similarity 41.5%; Pred. No. 0.0064;
Matches 22; Conservative 4; Mismatches 15; Indels 12; Gaps 3;

Qy 6 WMECSVSGDGIQRRTDCLGPOAQAPVPADFQHL---PKPTVYRGCWAGPC 55
Db 560 WGQCSVTGCGVAVRQRTCLG-----GVFGDHLCCQGPK-TEQRACDGGPC 603

Search completed: March 13, 2004, 07:46:08
Job time : 2.44688 secs
```